

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd

OM nucleic - nucleic search, using bw model

Run on: February 26, 2005, 17:00:07 ; Search time 2362 Seconds
(without alignments)
10216.214 Million cell updates/sec

Title: US-10-621-901-26
Perfect score: 498
Sequence: 1 ggccttgacaaagatgcct.....aaattctgcacaagagctgca 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
Searched:      4708233 beqs, 24227607955 residues
Total number of hits satisfying chosen parameters:  9416466
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%

Listing first 45 summaries

```

1:  GenEmb1:*
2:  gb_ba:*
3:  gb_htg:*
4:  gb_in:*
5:  gb_cm:*
6:  gb_ov:*
7:  gb_pat:*
8:  gb_ph:*
9:  gb_pl:*
10: gb_pr:*
11: gb_ro:*
12: gb_ats:*
13: gb_ay:*
14: gb_vi:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	189.4	38.0	3015	6	CQ578439	Sequence
2	189.4	38.0	3037	3	AY047504	Drosophil
3	125.4	25.2	6522	6	CQ578438	Sequence
4	125.4	25.2	17504	3	AC002473	Drosophil
5	125.4	25.2	88933	2	AC017870	Drosophil
6	125.4	25.2	171554	3	AC093019	Drosophil
7	125.4	25.2	288451	3	AE003456	Drosophil
8	84.8	17.0	250968	2	AC014140	Drosophil
9	83.2	16.7	2486	6	CQ583587	Sequence
10	83.2	16.7	6105	6	CQ583586	Sequence
11	83.2	16.7	154772	3	AC009214	Drosophil
12	83.2	16.7	165267	3	AC005711	Drosophil
13	83.2	16.7	246230	3	AE003636	Drosophil
14	72.4	14.5	5349	6	CQ578543	Sequence
15	65.8	13.2	2184	6	CQ578538	Sequence
16	65.8	13.2	4399	6	CQ578537	Sequence
17	59	11.8	2552	2	CQ578544	Sequence
18	43.6	8.8	175330	2	CR352286	Danio rerio
19	43.6	8.8	181320	5	BX547992	Zebrafish

20	42.2	8.5	222994	2	AC094289	Rattus no
21	41.8	8.4	2175	9	AF041643	Homo sapi
22	41.8	8.4	3258	9	AY273896	Homo sapi
23	41.8	8.4	5054	6	CQ851267	Sequence
24	41.8	8.4	5054	9	AK128854	Homo sapi
25	41.8	8.4	5238	9	HSMB07493	Homo sapi
26	41.8	8.2	3998	5	AJ719335	Homo sapi
27	40.8	8.2	185456	2	AC091379	Mus muscu
28	40.8	8.2	217530	10	AC121845	Mus muscu
29	40.4	8.1.1	2594	10	AF203701	Mus muscu
30	40.4	8.1	2675	10	BC071214	Mus muscu
31	40.4	8.1.1	2798	10	AF213560	Mus muscu
32	40.4	8.1.1	108609	2	AY379774	AY379774 Medicago
33	40.4	8.1.1	110000	8	CR381237_20	Continuation (21 o
34	40.4	8.1	158372	2	BX470144	Danio rerio
35	40.4	8.1.1	266607	2	BX571847	Danio rerio
36	40.2	8.1.1	8866	6	AX277977	Sequence
37	40.2	8.1.1	8866	6	AX333666	Sequence
38	40.2	8.1.1	16771	3	CEY32FEB	AL021475 Caenorhab
39	40	8.0	427939	3	U039993	U039993 Caenorhabdi
40	40	8.0	67575	3	AC099987	U0399987 Mus muscu
41	39.6	8.0	5882	6	AX345447	AX345447 Sequence
42	39.6	8.0	145685	2	AC015640	AC015640 Homo sapi
43	39.6	8.0	151939	9	AP001980	AP001980 Homo sapi
44	39.6	8.0	167960	2	AC027559	AC027559 Homo sapi
45	39.6	8.0	169801	2	AC018429	AC018429 Homo sapi

ALIGNMENTS

[illegible]

Db 1307 CAGTTATATCTGGAATGAGATGCTATCTACACGTTGGAGGTCCTACATGAC 1366
QY 241 GATTAACACCAAAAAATCTAAAAACACGACGACTGATTAATTTCTTATTTCTTCGTATG 300
Db 1367 GATTAATCTAAAAAATCTAAAACTCCGCGCTTTGTGATCTTTCATCTTTTGGCGATG 1426
QY 301 CTCGACCTGCAACCGGTTATGCTTGGCTAGCGCTGTCTCTAAAGTTTCAATTTGCCA 360
Db 1427 TTGGGCTCTGCGCATTTGGATAGCGCCCTGGCATCTTCTGCGCCCTGTATATGACCA 1486
QY 361 ACCTTGACCGCAACATTTGATTAATTAATGATCTAGATGTTTANGAGATGTTGTTAAGT 420
Db 1487 CAATTCATCTCGGTGATCAACAACAAGATCCCGCTGGCTGGCGCTGTGTGTGCGC 1546
QY 421 TGGGTGATATCTAGATCACTTAATATTTTGTGCAACGTTGATTTGATTTTCTTAA 480
Db 1547 TGGCTGTGATGCGCGCGCTGCTCTCTTCTCGGAGTTTCTCTCCATGTTTCCCAA 1606
QY 481 ATTCTGCCAAGAGCTG 496
Db 1607 GAATGCCCGACAGCAG 1622

RESULT 2
AY047504 3037 bp mRNA linear INV 16-AUG-2001
LOCUS Drosophila melanogaster GH01304 full length cDNA.
DEFINITION AY047504
ACCESSION AY047504.1 GI:15010375
KEYWORDS FLI-CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephyroidae; Drosophilidae; Drosophila.
1 (bases 1 to 3037)
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
Champe,M., Chavez,C., Dorsett,V., Fartan,D., Frise,E., George,R.,
Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
Nunoo,J., Paclebo,J., Paragas,V., Park,S., Phoumanavong,S., Wan,K.,
Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.
Direct Submission
Submitted (19-JUL-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.
Location/Qualifiers
1..3037
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/map="5BD1-5BD1"
/clone="GH01304"
1..3037

FEATURES
source
gene

/gene="CG3380"
/note="alignment with genomic scaffold AB003456"
/db_xref="FLYBASE:FBgn0034716"
800..2896
/gene="CG3380"
/note="longest ORF"
/codon_start=1
/product="GH01304p"
/protein_id="AAK7236.1"
/db_xref="GI:15010376"
/translation="M.PASGYSREPTDLPILKLTLLYIGVICISMTVAVFNGTI
TTBRKRPISKMTGIIISGNDISQTVSAVALIYAKGHRPRMIGPGLITVFCIL
TAPHPFLYGGEDALALISEFGMPDENATMEALIEORSKTLICRLNMGGAECVEGECN
FAPOLLFVAQFIISIGISLYTLTGVSVMDDTKSKTPALLSLYSFLMLGPAIGYA
LASFCILYIAQPMHPVINNKDPRMIGAMLMVLWGILSFSFGLSMFKPELPRAY
ARRKYENRREKRLSVKSTKERLALDOKTTEAASFODMKLTFRRLITNKTY
MONTLSSIFLYIGCTPYWYIETPKYIEVOROSAAISMTGTVALAFSAVGLISGFI
ISRYRPARIYMAAMVYIVGFLVAGILALAFICGRENESVIYNIHDSLAGNTTCN
SACSCDYVRYSFVCGENNTYISACIAGAKLLVNSBKRTFYDCISDDAGNITS
QPKRLTSPDLSNDEVSONTSVHSOLEALNMGAMPACVNCMTQFVAELAVWCILKF
VGASGRASNFIVSRCPKDKTAAAGFMTLCSMIAFIIPSPIFFGMVDRVCLVWGK
TCNKGNCWLYDPLASRYTLNFTAAVFIAGAIIFDLGVVYAKDLKIFEDVYEVEMK
IVQHEBEANENKTEI"

ORIGIN
Query Match 38.0%; Score 189.4; DB 3; Length 3037;
Best Local Similarity 61.3%; Pred. No. 1,8e-42;
Matches 304; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 1 GGCCCTGGAACAAGATCCCTTGACTCACCAGAAATATGAGCTGCTTACATCAAAAT 60
Db 1130 GGACCCGGGAGAGATCACCTTGCTGACCTCGGAGTTGGCGGAGATGCCGATGAAAT 1189
QY 61 GCCACATTTGAAGTTCTTAATAAAGAGAAACTTTGGCCACATGATGGCGAA 120
Db 1190 GCCACATGAGAGCTATTCAGAGACGACGATCCAAACCTATGCCGCTTGAATGAGAGA 1249
QY 121 GGAATTGAATGCGAAACAGAAAGGAGATCGACCTCAATAATTAATTTGTGCGC 180
Db 1250 GGAAGGAGAGGAGAGTGGTGAAGAACTTTGCCCCCACTTTTGCTCGTAGCG 1309
QY 181 CAATTAATTTCTGAGTCCGTGATCTTTATCTAACAATTTAGGTATCTTATGAGAC 240
Db 1310 CAGTTATATCTGGAATGAGATGCTATCTACACGTTGGAGATGTCCTACATGAGAC 1369
QY 241 GATTAACACCAAAAAATCTAAAAACACGACGACTGATTAATTTCTTATTTCTTCGTATG 300
Db 1370 GATTAATCTAAAAAATCTAAAACTCCGCTTTGTAGTCTTTCATCTTTTGGCGATG 1429
QY 301 CTCGACCTGCAACCGGTTATGCTTGGCTAGCGCTGTCTTAAAGTTTCAATTTGCCA 360
Db 1430 TTGGGCTCTGCGCATTTGGATAGCGCCCTGGCATCTTCTGCGCCCTGTATATGACCA 1489
QY 361 ACCTTGACCGCAACATTTGATTAATTAATGATCTAGATGTTTANGAGATGTTGTTAAGT 420
Db 1490 CAATTCATCTCGGTGATCAACAACAAGATCCCGCTGGCTGGCGCCCTGTGTGCGGCG 1549
QY 421 TGGGTGATATCTAGATCACTTAATATTTTGTGCAACGTTGATTTGATTTTCTTAA 480
Db 1550 TGGCTGTGATGCGCGCGCTGCTCTCTTCTCGGAGTTTCTCTCCATGTTTCCCAA 1609
QY 481 ATTCTGCCAAGAGCTG 496
Db 1610 GAATGCCCGACAGCAG 1625

RESULT 3
COS78438/c 6522 bp DNA linear PAT 02-FEB-2004
LOCUS COS78438
DEFINITION Sequence 6196 from Patent WO0171042.
ACCESSION COS78438
VERSION COS78438.1 GI:41640938

KEYWORDS	ORGANISM	REFERENCE	TITLE	JOURNAL	FEATURES	ORIGIN
Drosophila sp.	Drosophila sp.	Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.	Detection of 10,000 or more Drosophila genes and uses thereof	Patent: WO 0171042-A 6196 27-SEP-2001;	PE Corporation (NY) (US)	1
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.						
1						
location/Qualifiers						
1..6522						
/organism="Drosophila sp."						
/mol_type="unassigned DNA"						
/db_xref="taxon:7242"						
Query Match	25.2%	Score 125.4;	DB 6;	Length 6522;		
Best Local Similarity	65.6%	Pred. No. 2.4e-24;				
Matches 183;	Conservative 0;	Matches 96;	Indels 0;	Gaps 0;		
QY	1	GGCCCTGGACAGAAATGCCCTTCGATCTCACTCAACCGAATATGAGAGCTGTACAGATCAAAAT	60			
DB	3510	GGACCCGGCGAGATGCACTTGCTTGAACCTCGAGTTTCGGGGGATGCCGATGAAAT	3451			
QY	61	GCCACCATTTGAAGTTCTTAAATAAACAGAAAGAGAAATCTTTGTGCCACATGATGCCGA	120			
DB	3450	GCCACAAATGAGAGGTATGAGAGAGAGAGATCAACCAACCTATATGCGCTTGAATGAGAGA	3391			
QY	121	GGAGTTGAATGCGAAACAGAAAGAGAGCATCGCACTCAATATATATTTGTGCC	180			
DB	3390	GGAGGCGAGTCCGAGTGGTGAAGAAATCTTTGGCCCCCAATTTCTCTTCGTACGG	3331			
QY	181	CAATTAATTTCTGAGTGGTGGATCTTTACTACTACTTTAGAGTATCTTATATGAGAC	240			
DB	3330	CAGTTCATATCTGGAATGAGAGATGCGCTATATCTACAGCTTGGAGTGTCTCATATGAC	3271			
QY	241	GATTAACACCAAAATCTTAAACACACGACACTGATAGT	279			
DB	3270	GATTAATCAAAAAATCCAAAACTCGGCTTTGTGAGT	3232			
RESULT 4						
AC002473/c	71504 bp	DNA	linear	INV 20-JAN-1998		
DEFINITION	Drosophila melanogaster (PI DB02833 (D61))	DNA sequence, complete				
LOCUS	AC002473					
ACCESSION	AC002473	L81545	L81544	L81543	L81542	L81541
VERSION	AC000801	L81539	L81538	L81537	AC000797	AC000796
KEYWORDS	AC000799	AC0001823	AC001826	AC001824	AC001822	AC001337
ORGANISM	AC001825	AC001339				
REFERENCE	AC002473.1	GI:2337890				
AUTHORS	HTG.					
TITLE	Drosophila melanogaster (fruit fly)					
JOURNAL	Drosophila melanogaster					
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.					
AUTHORS	1 (bases 1 to 71504)					
TITLE	Ceintrix, S.E., Aghavan, A., Arcaina, T.T., Baxter, E., Doyle, C.M., Fafan, D.E., Flanagan, J., Houston, K.A., Hummachi, S.R., Kary, K., Kearney, L., Kim, S.H., Ko, C.L., Li, M., Lomocan, M.A., Mazda, P., Mok, M.S., Nixon, K., Pacleb, J.M., Park, S., Pfeiffer, B., Punch, D., Sanceo, R.E., Snit, E., Stevo, V., Subramanian, S., Towne, B., Wan, K.H., Whitelaw, K.R., Yee, A., Zhang, R., Zieran, L.T. and Kimmel, B.					
JOURNAL	Sequencing of Drosophila chromosome 2R, region 5B2					
REFERENCE	Unpublished (1997)					
AUTHORS	2 (bases 1 to 71504)					
TITLE	Martin, C.H., Alcivar, D.A., Arcaina, T.T., Bondoc, M.M., Chiang, A., Ciliz, P.A., Davis, C.A., Doyle, C.M., Ericsson, C.L., Fafan, D.E.,					

TITLE
 JOURNAL
 COMMENT
 Direct Submission
 Submitted (20-AUG-1997) Berkeley Drosophila Genome Project, MS 74-157, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US
 Sequence submitted by:
 Lawrence Drosophila Genome Project
 Berkeley Drosophila National Laboratory, MS 64-121 Berkeley, CA 94720
 For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www-hpc.lbl.gov/sequence-archive.html>) or send email to drosophila@genome.lbl.gov.
 Library location: 30-49.
 This pl was assembled from the following subclones: 2_a4 (L81545), 1_f7 (L81544), 2_e9 (L81543), 2_b2 (L81542), 2_h6 (L81541), 2_b11 (L81540), 1_f9 (AC000798), 2_d6, 2_g7 (AC000801), 2_c11 (L81539), 1_f5 (L81538), 2_f5 (L81537), 1_e12 (AC000797), 1_b7, 2_a11, 1_f11, 1_c2 (AC000796), 2_g8 (AC000802), 2_f2 (AC000800), 2_c4 (AC000799), 2_c9 (AC001823), 2_g9 (AC001826), 2_e3 (AC001824), 2_a1 (AC001822), 1_c11 (AC001337), 2_f6 (AC001338), 2_g3 (AC001825), 2_f8 (AC001339).
 FEATURES
 SOURCE
 location/Qualifiers
 1. 71504
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /chromosome="2R"
 /map="58C2-58C7"
 /clone="P1 D802833 (D61)"
 ORIGIN
 Query Match 25.2%; Score 125.4; DB 3; Length 71504;
 Best Local Similarity 65.6%; Pred. No. 2,5e-24;
 Matches 183; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
 QY 1 GGCCCTGGACAAGATGCGCTTGCATCTACCCAGATATGAGAGCTGTACGATCAAAAT 60
 Db 61567 GGACCGGCGCAGAGATGCATCTGCTCGAAGCTTGGGGAGTCCGATGAATAAT 61508
 QY 61 GCCACCACTTGAATCTCTTAATAACAGAGAGGAAACTTTGGCCACACTGATGGCGAA 120
 Db 61507 GCCACCAATGAGGCTTTCGAGGAGCGACGATCCAAACCTTATGCCGTTGAATGAGGA 61448
 QY 121 GGAGTTGAATGCGAAACAGAGAGGAGAGCATGCACTCAATATATATCTATTTGCTGCC 180
 Db 61447 GGAGCGGAGTGCAGAGGTGCGTGAAGAACTTTGGCCCCCACTTTTGTCTTCGTAAGG 61388
 QY 181 CAATTAATTTCTGAGTGGTGGATCTTTATATACATCTTAAGGTATCTATATGAC 240
 Db 61387 CAGTTCAATCTTGAATTTGGAGGATGCTATATCTACACGTTGGGAGTGTCTTACATGAGAC 61328
 QY 241 GATTAACCCAAAATCTTAAACACGACGATGATAGT 279
 Db 61327 GATTAATCTTAAATAATCCAAAATCTCGGCTTTGTTGAGT 61289
 RESULT 5
 AC017870/c
 LOCUS AC017870 88933 bp DNA linear HTG 09-DEC-1999
 DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.
 ACCESSION AC017870
 VERSION AC017870.1 GI:6553320
 KEYWORDS HTG; HTGS PHASE2.
 SOURCE
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 88933)
AUTHORS
Adams,M. and Venter,J.C.
JOURNAL
Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
This sequence was identified as CDM:10212423 by the submitter. For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. .88933
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

ORIGIN

Query Match 25.2%; Score 125.4; DB 2; Length 88933;
Best Local Similarity 65.6%; Pred. No. 2.5e-24;
Matches 183; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 1 GGGCCCTGGACAAGATGCCCTTTCAGCTCACCACGATATGAGCTGTTCATGATCAAAAT 60
DB 64160 GGACCCCGCGAGATGATGACTTGTCTGACCTCGGAGTTCGCGGAGATGCCCGAATGAAAAAT 64101
QY 61 GCCACCATTAAGTCTTAATAAACAAGAGAAACCTTTGGCCACACTGATGGCGAA 120
DB 64100 GCCACATGAGAGGCTATCGAGAGACAGCATCCAAAGCCCTTATGCCCTGAATGGAGGA 64041
QY 121 GGAATTGATATGCGAAGACAGAAAGGAGCATGCGACCTCAATTAATATTGCTGCC 180
DB 64040 GGAGCGGAGTGGCAGAGTCCGATGGAAGAACTTTGCGGCCCAACTTTGCTTCTGTAGCG 63981
QY 181 CAATTAATTTCTGGAGTGGTGTATCTTTATCTACACTTTAGTGTATCTTATATGAC 240
DB 63980 CAGTTCTATATCTGGAATTTGAGAGATCGCTATCTACACAGTTGGAGATGCTCTCATGAC 63921
QY 241 GATPACACCAAAAATCTTAACACCGACGACTGATAGT 279
DB 63920 GATPAACTAATAAAATCAAAATCCGCGCTTTGTAGT 63882

RESULT 6
AC099019 171594 bp DNA linear INV 08-NOV-2001
LOCUS
DEFINITION
Drosophila melanogaster, chromosome 2R, region 58C-58D, BAC clone
BACR02012, complete sequence.
AC099019
AC099019.1 GI:16798951
HTG.
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 171594)
Celisner,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Branton,R.C., Rogers,Y., An,H., Baldwin,D., Bazon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champs,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferreira,S., Frise,E., Galle,R.F., Gary,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Scapleton,M., Strong,R., Svitskas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Sequencing of Drosophila chromosome, 2R, region 58C-58D

TITLE

JOURNAL
REFERENCE
Unpublished
2 (bases 1 to 171594)
AUTHORS
Celisner,S.E., Adams,M.D., Krommiller,B., Tyler,D., Wan,K.H., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Branton,R.C., Rogers,Y., An,H., Baldwin,D., Bazon,J., Beeson,K.Y., Busam,D.A., Carlson,J.W., Center,A., Champs,M., Davenport,L.B., Dietz,S.M., Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D., Ferreira,S., Frise,E., Galle,R.F., Gary,N.S., George,R.A., Gonzalez,M., Houck,J., Hoskins,R.A., Hostin,D., Howland,T.J., Ibegwam,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A., McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B., Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F., Scapleton,M., Strong,R., Svitskas,R., Tector,C., Williams,S.M., Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C.
Direct Submission
Submitted (08-NOV-2001) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.

FEATURES
source
1. .171594
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2R"
/map="58C-58D"
/map="BACR02012 (D1312)"
/clone_id="RPCT-98 (Roswell Park Cancer Institute Drosophila melanogaster BAC library, partial EcoRI in pBACs.6)"

ORIGIN

Query Match 25.2%; Score 125.4; DB 3; Length 171594;
Best Local Similarity 65.6%; Pred. No. 2.5e-24;
Matches 183; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 1 GGGCCCTGGACAAGATGCCCTTTCAGCTCACCACGATATGAGCTGTTCATGATCAAAAT 60
DB 20134 GGACCCCGCGAGATGATGACTTGTCTGACCTCGGAGTTCGCGGAGATGCCCGAATGAAAAAT 20193
QY 61 GCCACCATTAAGTCTTAATAAACAAGAGAAACCTTTGGCCACACTGATGGCGAA 120
DB 20194 GCCACATGAGAGCTATCGAGAGACAGCATCCAAAGCCCTTATGCCCTGAATGGAGGA 20253
QY 121 GGAATTGATATGCGAAGACAGAAAGGAGCATGCGACCTCAATTAATTAATATTGCTGCC 180
DB 20254 GGAGCGGAGTGGCAGAGTCCGATGGAAGAACTTTGCGGCCCAACTTTGCTTCTGTAGCG 20313
QY 181 CAATTAATTTCTGGAGTGGTGTATCTTTATCTACACTTTAGTGTATCTTATATGAC 240
DB 20314 CAGTTCTATATCTGGAATTTGAGAGATGCTATCTACACGTTGGAGATGCTCTCATGAC 20373
QY 241 GATPACACCAAAAATCTTAACACCGACGACTGATAGT 279
DB 20374 GATPAACTAATAAAATCAAAATCCGCGCTTTGTAGT 20412

RESULT 7
AE003456/c 288451 bp DNA linear INV 22-MAR-2004
LOCUS
DEFINITION
Drosophila melanogaster chromosome 2R, section 64 of 74 of the complete sequence.

ACCESSION	AE003456 AE002575 AE013599	JOURNAL	Genome Biol. 3 (12), RESEARCH0083 (2002)
VERSION	AE003456.3 GI:45445355	MEDLINE	22426069
KEYWORDS		PUBMED	12537572
SOURCE		REFERENCE	4 (bases 1 to 288451)
ORGANISM	Drosophila melanogaster (fruit fly)	AUTHORS	Kaminker, J.S., Bergman, C.M., Kronmiller, B., Carlson, J., Svitskas, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E., Rubin, G.M., Ashburner, M. and Celniker, S.E.
REFERENCE	Drosophila melanogaster (fruit fly) Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 288451)	TITLE	The transposable elements of the <i>Drosophila melanogaster</i> euchromatin: a genomics perspective
AUTHORS	Adam, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Gall, R.F., George, R.A., Lewis, S.E., Richards, S., Ashburner, M., Henderson, S.N., Sutton, G.G., Wortman, J.R., Venter, L., White, M., Zheng, X., Smith, C.J., Rong, Y., H. Blakes, R.G., Rogers, Y.H., Blakes, R.G., Chapple, M., Pfeiffer, B.D., Man, K.H., Doyle, C., Baxter, E.G., Helt, G., Nelson, C.R., Gaber, D., Abril, J.F., Agbayani, A., An, H.J., Andrews-Pfannkoch, C., Baldwin, D., Ball, W.R., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M., Beeson, K.Y., Benos, P.V., Bernier, B.P., Bhandari, D., Bolintsov, S., Borkov, D., Botchan, M.R., Bouck, J., Brokstein, P., Brotler, P., Burtis, K.C., Busam, D.A., Butler, H., Caden, E., Center, A., Chandra, I., Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B., Davies, P., de Pablo, B., Delcher, A., Deng, Z., Mays, A.D., Dew, I., Dietz, S.M., Dodson, K., Dou, P., Downes, M., Dugan-Rocha, S., Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferrar, C., Ferrier, S., Fleischmann, W., Foster, C., Gabriellian, A.E., Garg, N.S., Gelbart, W.M., Glasser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z., Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J., Hernandez, J.R., Houck, J., Hostin, D., Houston, K.A., Howland, T.J., Wei, M.H., Idagawa, C., Jaisi, M., Kalish, F., Karpen, G.H., Ke, Z., Kennison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Laoko, P., Lei, Y., Levitsky, A.A., Li, J., Li, Z., Liang, Y., Lin, X., Lin, X., Mathe, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C., Moritz, J., Mosher, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L., Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusskern, D.R., Pacle, J.M., Palazzolo, M., Peltman, G.S., Pan, S., Pollard, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I., Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, E., Svitskas, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A., Weinstock, G.M., Weissenbach, J., Williams, S.M., Woodgett, J., Worley, K.C., Wu, D., Yang, S., Yao, Q., Ye, J., Yeh, R.F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.	TITLE	The genome sequence of <i>Drosophila melanogaster</i>
JOURNAL	Science 287 (5461), 2185-2195 (2000)	COMMENT	Submitted (10-MAR-2004) FlyBase, Harvard University, Biological Laboratories, 16 Divinity Avenue, Cambridge, MA 02138, USA
MEDLINE	20196006	FEATURES	On Mar 15, 2004 this sequence version replaced gi:21626512.
PUBMED	10731132	source	Location/Qualifiers
REFERENCE	2 (bases 1 to 288451)	gene	1. 288451
AUTHORS	Celniker, S.E., Wheeler, D.A., Kronmiller, B., Carlson, J.W., Halpern, A., Patel, S., Adams, M., Dugan, S.P., Frise, E., Hodgson, A., George, R.A., Hoskins, R.A., Lavey, T., Muzny, D.M., Nelson, C.R., Pacle, J.M., Park, S., Pfeiffer, B.D., Richards, S., Seligson, E.J., Svitskas, R., Taber, P.E., Man, K., Stapleton, M., Sutton, G.G., Venter, C., Weinstock, G., Scherer, S.E., Myers, E.W., Gibbs, R.A. and Rubin, G.M.	mRNA	/db_xref="FLYBASE:FBgn0034680" complement(join(6807..7177,39294..39621,41306..41460)) /locus_tag="CG13499" /map="58B1-58B1"
TITLE	Finishing a whole-genome shotgun: release 3 of the <i>Drosophila melanogaster</i> euchromatic genome sequence	mRNA	/db_xref="FLYBASE:FBgn0034680" complement(join(6807..7177,39294..39621,41306..41460)) /locus_tag="CG13499" /codon_start=1 /product="CG13499-PB" /protein_id="AAFA6785.2" /db_xref="GI:21626513"
JOURNAL	Genome Biol. 3 (12), RESEARCH0079 (2002)	mRNA	/db_xref="FLYBASE:FBgn0034680" complement(join(6807..7177,39294..39621,41306..41460)) /locus_tag="CG13499" /product="CG13499-RD" /locus_tag="CG13499"
MEDLINE	22426065	mRNA	/db_xref="FLYBASE:FBgn0034680" complement(join(6807..7177,39294..39621,41306..41460)) /locus_tag="CG13499" /product="CG13499-RC" /locus_tag="CG13499"
PUBMED	12537568	mRNA	/db_xref="FLYBASE:FBgn0034680" complement(join(6807..7177,39294..39621,41306..41460)) /locus_tag="CG13499" /product="CG13499-PA" /protein_id="AAFA68195.1"
REFERENCE	3 (bases 1 to 288451)	mRNA	/db_xref="GI:21626514"
AUTHORS	Misra, S., Crosby, M.A., Mungall, C.J., Matthews, B.B., Campbell, K.S., Hradecsky, P., Huang, Y., Kaminker, J.S., Milburn, G.H., Prochuk, S.E., Smith, C.J., Tupy, J.L., Whitfield, E.J., Bayraktaroglu, L., Berman, B.P., Bettencourt, B.R., Celniker, S.E., de Grey, A.D., Drysdale, R.A., Harris, N.L., Richter, J., Russo, S., Schroeder, A.J., Shu, S.Q., Stapleton, M., Yamada, C., Ashburner, M., Gelbart, W.M., Rubin, G.M. and Lewis, S.E.	CDS	/db_xref="GI:21626514"
TITLE	Annotation of the <i>Drosophila melanogaster</i> euchromatic genome: a systematic review	CDS	

Query Match	Best Local Similarity	Score	IDB	Length
Matches 183; Conservative	65.6%; Pred. No. 2.6e-24; Mismatches 96; Indels 0; Gaps 0;	125.4; DB 3;	288451;	
QY 1	GGCCCTGGACAAGATGCCCTTGCAGACTCACCACGAAATATGAGCTGTCTACATCAAAAT	60		
Db 228092	GGACCCGGGAGAGATGCACTTGCTGCACTGCGAGATTTGGCGGAGATGCCAGATGAAAAAT	228033		
QY 61	GCCACCATTTGAAGTCTTATATAAACAGAAAGAGAAAACTTTGTGCCACACTGATGCCGAA	120		
Db 228032	GCCACAAATGAGAGCTATCGAGAGAGCGATCCACAAAGACCTTATGCGCTTGAATGAGGGA	227973		
QY 121	GGAGTTGAATGGGAAACAGAAAGGAGACATTCGACCTCAATAATATCTATTGCTGCC	180		
Db 227972	GGAGCGGAGTGGAGGTGCGTGAAGGAACTTTGCCGCCCAACTTTTGCTTTCGATGCG	227913		
QY 181	CAATTAATTTTCGAGATCGGTGATTTTATATCTACACTTTAGTGTGATTCCTATATGAC	240		
Db 227912	CAGTTCAATCTCGGAATTTGAGAGATGCTATACTACACGTTGGGAGATGTCTCAATGAC	227853		
QY 241	GATTAACACCAAAAAATCTTAAACACCAAGACGATGATAGT	279		

Db	227852	GATAA	TA	CTA	AAAAATCCAA	AACTCCGCGTTTG	TGAGT	227814		
RESULT 8										
AC014140/c										
DEFINITION	AC014140					250968 bp	DNA	linear	HTG 16-NOV-1999	
LOCUS	Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***.									
ACCESSION	AC014140									
VERSION	AC014140.1					GI:6437195				
KEYWORDS	HTG; PHASE2.									
SOURCE	Drosophila melanogaster (fruit fly)									
ORGANISM	Drosophila melanogaster									
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.									
REFERENCE	1 (bases 1 to 250968)									
AUTHORS	Adams, M. and Venter, J.C.									
TITLE	Direct Submission									
JOURNAL	Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA									
COMMENT	This sequence was identified as CDM:1023167 by the submitter. For further information on this sequence e-mail to fly@celera.com.									
	* NOTE: This is a 'working draft' sequence.									
	* This sequence will be replaced									
	* by the finished sequence as soon as it is available and									
	* the accession number will be preserved.									
FEATURES										
source	1..250968									
	location/Qualifiers									
	/organism="Drosophila melanogaster"									
	/mol_type="genomic DNA"									
	/db_xref="taxon:7227"									
ORIGIN										
Query Match	17.0%;					Score 84.8;				
Best Local Similarity	53.9%;					Pred. No. 7.9e-13;				
Matches 173;	Conservative					0;				
	Mismatches 148;					Indels				
	0;					Gaps				
	0;									
Qy	173	TTGCTGCCCA	TTAA	TTTCTG	GAGTGGTGATCTTTAT	ACTACACTT	AGTGTAT	CCCT	232	
Db	205056	TTTCTCTC	CAAGTTC	GTTC	TGCGGCGTGCGACACCTTG	TA	CTACCTCC	TGGGACAGACCT	204997	
Qy	223	ATATGAC	AGATTA	ACCCAA	AAAAATTTAA	ACCCAGCACTG	ATTAAGTT	TTTCTTAT	292	
Db	204996	ACCTCGA	GACAA	CAACCA	AGAAAGAA	CACATCCGCTC	ATGCGCGTGCC	AGTGCAC	204937	
Qy	293	TTCGAT	CTCGGAC	CTGCAAC	CGGTATAGCCTTG	AGCTGTCTG	TTAAAGTT	TTTACA	352	
Db	204936	TCAGAT	ATATG	ACCACT	AGTCGTCGGGATCTTTT	GGTTTCA	TATGCTT	TAACCTTCA	204877	
Qy	353	TTTGGC	CAAGCTT	GACG	CCCAACATTTG	ATATATGATCT	ATGATGT	TANAGAT	412	
Db	204876	TCGAT	CCAA	CAAGAC	CGCGCTTG	ATCGACAG	AGACCCCGCG	TGAGGCTCGT	204817	
Qy	413	GGTTA	AGTTGG	GTGAT	ACTAGATCA	CTTAAATATTTT	TGCAACGTT	GATTTGAT	472	
Db	204816	GCGTC	GCGGTG	GTATCT	CTGGGCA	CCCTCATGTG	CGCTTTC	CGGCTTAT	CGACTGT	204757
Qy	473	TTCT	AAAAAT	CTGCCA	AG	493				
Db	204756	TTCC	CAAGCA	ATTGCC	CAAG	204736				
RESULT 9										
LOCUS	COS83587					2486 bp	DNA	linear	PAT 02-FEB-2004	
DEFINITION	Sequence 11345 from Patent WO0171042.									
ACCESSION	COS83587									
VERSION	COS83587.1					GI:41644371				
KEYWORDS										
SOURCE	Drosophila sp.									
ORGANISM	Drosophila sp.									
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;									

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS Venter, J.C., Adams, M., Li, P.W., and Myers, E.W.
TITLE Detection kits, such as nucleic acid arrays, for detecting the
JOURNAL expression of 10,000 or more Drosophila genes and uses thereof
Patent: WO 0171042-A, 11345 27-SEP-2001;
PE Corporation (NY) (US)

FEATURES
source Location/Qualifiers
1..2486
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"

ORIGIN

Query Match 16.7%; Score 83.2; DB 6; Length 2486;
Best Local Similarity 54.0%; Pred. No. 2e-12;
Matches 169; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

181 CAATTATTTCTGGAGTCGGTGGATCTTTATCTAGTATCTATATAGAC 240
186 CAATTCCTTCTGGAGTCGGGACACCTTGTACTCTCTGGACAGACCTTACTGAC 875
241 GATTAACACCAAAATCTAAACACAGACCTGATTAAGTTTCTTCTTCTGATG 300
876 GACAAACCAAGAGACGACATCCGCTATGCTGGCGGTCATGACAGATG 935
301 CTGGAGCTGCAACCGCTTATGCTTGGCTAGCTGTCTTAAAGTTTCTTCTGATG 360
336 ATTGACACCATGCTGGATCTTTTCTGTTCTATGCTTAAACCTTCTATGATCA 995
361 ACCTGACGCAACATGATTAATGATGCTGATGATGATGATGATGATGATGATGAT 420
996 ACCAAGACGCTGTTGATGACAGACGACACCCCGCTGCTAGTGGCTGCTGCTGCTG 1055
421 TGGGTGATCTAGGATCACTTAAATTTTCTGCAAGTTGATGATGATGATGATGATGAT 480
1056 TGGGTGATCTGCGGACCTCATGCTGCTTCTGCTGCTTATGCTGCTGCTGCTGCTG 1115
481 ATTCTGCCAAGAG 493
1116 CAATTGCCAAG 1128

RESULT 10
CQ583586 6105 bp DNA linear PAT 02-FEB-2004

LOCUS CQ583586
DEFINITION Sequence 11344 from Patent WO0171042.
ACCESSION CQ583586
VERSION CQ583586.1 GI:41644370
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS Venter, J.C., Adams, M., Li, P.W., and Myers, E.W.
TITLE Detection kits, such as nucleic acid arrays, for detecting the
JOURNAL expression of 10,000 or more Drosophila genes and uses thereof
Patent: WO 0171042-A, 11344 27-SEP-2001;
PE Corporation (NY) (US)

FEATURES
source Location/Qualifiers
1..6105
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"

ORIGIN

Query Match 16.7%; Score 83.2; DB 6; Length 6105;
Best Local Similarity 52.8%; Pred. No. 2.1e-12;
Matches 178; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

157 CTTCAATTAATTAATTTGCTGCCAATTAATTTGAGTCGGTGGATCTTATACTAC 216
3221 CCACTTGCTCTCATCTTCTCTGCGAGTTGCTTCTGGGCGTGGCAACACCTTGTACTAC 3280

217 ACTTTAGGTATCTTATATGACGATTAACCAAAATCTTAAACACAGACTGATA 276
3281 TCCCTGGAGACGACCTTACCTGACGACACCAAGAGAGACATCTGCTCATGCTG 3340

277 AGTTTCTTATTTCTTCTGATGCTGACCTGCAACCGGTTATGCTTGGCTAGCGTC 336
3341 GCGGTGCGCATGCGACTCAGATGATGATGACAGCTGCTGGATCTTTTCGTTTATA 3400

337 TGTCTAAGTTCTTACATTTTGGCCAGCTGACGCCCAACATGATTAATATGATCTAGA 336
3401 TCGCTTAACACCTTATCATGATTCACCAAGACCGCTGATGACAGACAGACCCCGC 3460

397 TGGTTAGACATGCTGCTTAAAGTTGGGTATATCTAGATCACTTAAATTTTTCGA 456
3461 TGGCTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3520

457 ACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 493
3521 GGTCTTATGCGACTGTTTCCCAAGCAATTCGCCAAG 3557

RESULT 11
AC009214 154772 bp DNA linear INV 31-JUL-2004

LOCUS AC009214
DEFINITION Drosophila melanogaster clone BACR05E09, complete sequence.
ACCESSION AC009214
VERSION AC009214.8 GI:50872371
KEYWORDS
SOURCE HTG.
ORGANISM Drosophila melanogaster (fruit fly)

REFERENCE
AUTHORS Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 154772)
Celinker, S.E., Asghar, A., Arcania, T.T., Baxter, E., Blazek, R.G.,
Butenhoff, C., Champagne, M., Chavez, C., Chew, M., Ciesiolka, L.,
Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
Hoskins, R.A., Houston, K.A., Hummasti, S.R., Katta, K., Kearney, L.,
Kim, E., Lee, B., Lewis, S., Li, P., Lomutan, M.A., Mazda, P.,
Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleb, J.M., Park, S.,
Pfeiffer, B., Poon, L., Seguela, A., Sethi, H., Smit, E.,
Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zierman, L.L. and
Rubin, G.M.

TITLE Direct Submission
JOURNAL Direct Submission
Submitted (06-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
2 (bases 1 to 154772)

REFERENCE
AUTHORS Celinker, S., Carlson, J., Wan, K., Pfeiffer, B., Frise, E., George, R.,
Hoskins, R., Stapleton, M., Pacleb, J., Park, S., Svirskas, R.,
Smith, E., Yu, C. and Rubin, G.

TITLE Direct Submission
JOURNAL Submitted (31-JUL-2004) Berkeley Drosophila Genome Project, MS
64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, US

COMMENT
On Jul 31, 2004 this sequence version replaced gi:15451504.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This sequence submission incorporates changes made during
revelation of the assembly or fingerprint verification of the
clone. For further information about this sequence, including its
location and relationship to other sequences, please visit our
sequence archive web site (<http://www.fruitfly.org/sequence/>) or
send email to bdg@fruitfly.org.

FEATURES
source Location/Qualifiers
1..154772
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"

/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2L"
/map="33E-33F"
/clone="BAC clone BACR05E09 (D865)"
/clone_11b="RPCT-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBAC3.6)"

ORIGIN

Query Match 16.7%; Score 83.2; DB 3; Length 154772;
Best Local Similarity 52.8%; Pred. No. 2.2e-12;
Matches 178; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

QY 157 CCTCAATAATACTATTGCTGCGCAATTAATTTCTGAGTCGGTGATCTTTACTAC 216
DB 140248 CCACTTGCTGCTACTTTCTCTGCACTTGCTTGCGCGTGCGCAACCTTGATAC 140307
QY 217 ACTTTAGGTATCTTATATGACGATTAACCAAAAATCTAAACACGACACTGATA 276
DB 140308 TCCCTGGGACAGACTACCTGACGACCAACCAAGACGAACACTCCGCTCATGCTG 140367
QY 277 AGTTTCTTATTTCTTCTGATGCTGACCTGCAACCGGTTATGCTTGCTAGGCTC 336
DB 140368 GCGGTGGCCATGCGACTCAGAGATGATGACAGTGGGATTTCTTTTCTGTTTCTATA 140427
QY 337 TGTCTAAGTTCTACATTTGCGCCACGTTGACGCCAACATTTGATATATGATCTAGA 396
DB 140428 TCGCTTAAACCTTCACTGATCCAGCAACAGCCGTTGATGACAGCAAGACCCCGC 140487
QY 397 TGGTTANAGCATGCTGTTAAGTTGGGTATACAGATCAACTTAATATTTTGTGA 456
DB 140488 TGGCTAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 140547
QY 457 ACGTTGATGGATTTGTTCTCTAAATTCGCAAG 493
DB 140548 GGTCTTATCGAGCTGTTCCCAAGCAATTGCCCAAG 140584

RESULT 12
AC005711/c 165267 bp DNA linear INV 31-JUL-2004
LOCUS AC005711
DEFINITION Drosophila melanogaster clone BACR48M09, complete sequence.
AC005711.2 GI:50872348
VERSION
KEYWORDS HTG.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 165267)

REFERENCE
AUTHORS Celniker,S.E., George,R.A., Galle,R., Svitskas,R.R., Hoskins,R.A.,
Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Chavez,C.,
Chew,M., Doyle,C.M., Farfan,D.E., Flanagan,J., Houston,K.A.,
Hummer,S.R., Karrera,K., Kearney,L., Kim,S.H., Lee,B.,
Lomocan,M.A., Mak,J., Mazda,P., Mok,M.S., Moshrefi,A.R.,
Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Punch,E.,
Smit,E., Twomey,B., Wan,K.H., Whiteley,K.R., Yee,A., Zhang,R.,
Zieren,L.L. and Kimmel,B.E.

TITLE Direct Submission
JOURNAL Submitted (26-SEP-1998) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
REFERENCE 2 (bases 1 to 165267)
AUTHORS Celniker,S., Carlson,J., Wan,K., Pfeiffer,B., Frise,E., George,R.,
Hoskins,R., Stapleton,M., Pacleb,J., Park,S., Svitskas,R.,
Smit,E., Yu,C. and Rubin,G.

TITLE Direct Submission
JOURNAL Submitted (31-JUL-2004) Berkeley Drosophila Genome Project, MS
64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720 US
COMMENT On Jul 31, 2004 this sequence version replaced gi:4056405.
Sequence submitted by:

Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This sequence submission incorporates changes made during
revaluation of the assembly or fingerprint verification of the
clone. For further information about this sequence, including its
location and relationship to other sequences, please visit our
sequence archive web site (<http://www.fruitfly.org/sequence/>) or
send email to bdgp@fruitfly.org.

FEATURES

source

1. 165267
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2L"
/map="33D3-33E8"
/clone="BAC clone BACR48M09 (D433)"
/clone_11b="RPCT-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBAC3.6)"

ORIGIN

Query Match 16.7%; Score 83.2; DB 3; Length 165267;
Best Local Similarity 52.8%; Pred. No. 2.2e-12;
Matches 178; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

QY 157 CCTCAATAATACTATTGCTGCGCAATTAATTTCTGAGTCGGTGATCTTTACTAC 216
DB 143273 CCACTTGCTGCTACTTTCTCTGCACTTGCTTGCGCGTGCGCAACCTTGATAC 143214
QY 217 ACTTTAGGTATCTTATATGACGATTAACCAAAAATCTAAACACGACACTGATA 276
DB 143213 TCCCTGGGACAGACTACCTGACGACCAACCAAGACGAACACTCCGCTCATGCTG 143154
QY 277 AGTTTCTTATTTCTTCTGATGCTGCAACCGGTTATGCTTGCTAGGCTC 336
DB 143153 GCGGTGGCCATGCGACTCAGAGATGATGACAGTGGGATTTCTTTTCTGTTTCTATA 143094
QY 337 TGTCTAAGTTCTACATTTGCGCCACGTTGACGCCAACATTTGATATATGATCTAGA 396
DB 143093 TCGCTTAAACCTTCACTGATCCAGCAACAGCCGTTGATGACAGCAAGACCCCGC 143014
QY 397 TGGTTANAGCATGCTGTTAAGTTGGGTATACAGATCAACTTAATATTTTGTGA 456
DB 143013 TGGCTAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 142974
QY 457 ACGTTGATGGATTTGTTCTCTAAATTCGCAAG 493
DB 142973 GGTCTTATCGAGCTGTTCCCAAGCAATTGCCCAAG 142917

RESULT 13

AE003636/c 246230 bp DNA linear INV 15-MAR-2004
LOCUS AE003636
DEFINITION Drosophila melanogaster chromosome 2L, section 45 of 83 of the
complete sequence.
AE003636 AE0014134
ACCESSION AE003636 GI:22946331
VERSION
KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

1 (bases 1 to 246230)
Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
Brandon,R.C., Rogers,Y.H., Blazej,R.G., Chame,M., Pfeiffer,B.D.,
Wan,K.H., Doyle,C., Baxter,E.G., Heit,G., Neilson,C.R., Gabot,G.L.,
Abril,J.F., Agbayani,A., An,H.J., Andrews-Plannkocn,C., Baldwin,D.,

TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Beslow, R.M., Baer, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M. Beeson, K.Y., Benos, P.V., Berman, B.P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M.R., Bouck, J., Brokstein, P., Brottier, P., Buttle, K.C., Cheung, D.A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J.M., Cawley, S., Dahlke, C., Davey, L.B., Davies, P., de Pablo, B., Delcher, A., Deng, Z., May, A.D., Dew, I., Dieter, S.M., Dodson, K., Dou, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferreira, S., Fleischmann, W., Foster, C., Gabriellian, A.E., Garg, N.S., Gelbart, W.M., Glasser, K., Glodde, A.G., Gong, F., Gottlieb, J.H., Gu, Z., Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J., Hernandez, J.R., Houck, J., Hoehn, D., Houston, C., Howland, T.J., Wei, M.H., Ibbagman, C., Jalali, M., Kalush, F., Karpen, G.H., Ke, Z., Kemison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C., Kraevitz, S., Kuip, D., Lai, Z., Lasco, P., Lei, Y., Levitsky, A.A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Mekulov, G., Milshina, N.V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L., Murry, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nuskern, D.R., Pacleb, J.M., Palazolo, M., Pittman, G.S., Pan, S., Pollard, J., Pu, V., Reese, M.G., Rehner, K., Remington, K., Saunders, R.D., Scheeler, F., Shen, H., Shie, B.K., Sidon-Kiamos, I., Simpton, M., Skupel, M.P., Smith, T., Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, E., Svitskas, R., Tecor, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Weasman, D.A., Weinstock, G.M., Weissbach, J., Williams, J.M., Woodger, T., Woley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, Y., Yeh, R.F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.
TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	The genome sequence of <i>Drosophila melanogaster</i> <i>Science</i> 287 (5461), 2185-2195 (2000)
TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	(bases 1 to 246230) Celinker, S.E., Wheeler, D.A., Krommiller, B., Carlson, J.W., Halpern, A., Patel, S., Adams, M., Champe, M., Dugan, S.P., Frise, E., Hedgson, A., George, R.A., Hoskins, R.A., Laverty, T., Murthy, D.M., Neilson, C.R., Pacleb, J.M., Park, S., Pfeiffer, B.D., Richards, S., Sodergren, E.J., Svitskas, R., Taber, P.E., Wan, K., Stapleton, M., Sutton, G.G., Venter, C., Weinstock, G., Scherer, S.E., Myers, E.W., Gibbs, R.A. and Rubin, G.M.
TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Finishing a whole-genome shotgun: release 3 of the <i>Drosophila</i> <i>melanogaster</i> euchromatic genome sequence <i>Genome Biol.</i> 3 (12), RESEARCH0079 (2002)
TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	(bases 1 to 246230) Mitra, S., Crosby, M.A., Mungall, C.J., Matthews, B.B., Campbell, K.S., Hardcastle, P., Huang, Y., Kaminker, J.S., Millburn, G.H., Prochnik, S.E., Smith, C.D., Tu, J.Y., Whitfield, E.J., Bayraktarov, L., Berman, B.P., Betencourt, B.R., Celinker, S.E., de Grey, A.D., Drysdale, R.A., Harris, N.L., Richter, J., Russo, S., Schroeder, A.J., Shu, S.Q., Stapleton, M., Yamada, C., Ashburner, M., Gelbart, W.M., Rubin, G.M. and Lewis, S.E.
TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Annotation of the <i>Drosophila melanogaster</i> euchromatic genome: a systematic review <i>Genome Biol.</i> 3 (12), RESEARCH0083 (2002)
TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	(bases 1 to 246230) Kaminker, J.S., Bergman, C.M., Kronmiller, B., Carlson, J., Svitskas, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E., Rubin, G.M., Ashburner, M. and Celinker, S.E.
TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	The transposable elements of the <i>Drosophila melanogaster</i> euchromatin: a genomics perspective <i>Genome Biol.</i> 3 (12), RESEARCH0084 (2002)
TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	(bases 1 to 246230) Adams, M.D., Celinker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.
TITLE	JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	Submitted (21-MAR-2000) <i>Celera Genomics</i> , 45 West Gude Drive,

REFERENCE	REFERENCE	REFERENCE	REFERENCE	REFERENCE
AUTHORS	Rockville, MD 20850, USA	CONSTRM	6 (bases 1 to 246230)	JOURNAL
TITLE	FlYBase	TITLE	Direct Submission	JOURNAL
JOURNAL	Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA	JOURNAL	Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA	JOURNAL
REFERENCE	7 (bases 1 to 246230)	REFERENCE	7 (bases 1 to 246230)	REFERENCE
AUTHORS	FlYBase	CONSTRM	Direct Submission	JOURNAL
TITLE	Submitted (10-MAR-2004) FlYBase, Harvard University, Biological Laboratories, 16 Divinity Avenue, Cambridge, MA 02138, USA	TITLE	Submitted (10-MAR-2004) FlYBase, Harvard University, Biological Laboratories, 16 Divinity Avenue, Cambridge, MA 02138, USA	JOURNAL
JOURNAL	On Sep 16, 2002 this sequence version replaced gi:10728739.	JOURNAL	On Sep 16, 2002 this sequence version replaced gi:10728739.	JOURNAL
REFERENCE	Location/Qualifiers	REFERENCE	Location/Qualifiers	REFERENCE
FEATURES	1. .246230	FEATURES	1. .246230	FEATURES
SOURCE	/organism="Drosophila melanogaster"	SOURCE	/organism="Drosophila melanogaster"	SOURCE
gene	/mol_type="genomic DNA"	gene	/mol_type="genomic DNA"	gene
gene	/db_xref="taxon:7227"	gene	/db_xref="taxon:7227"	gene
gene	/chromosome="2L"	gene	/chromosome="2L"	gene
gene	<809. .>4062	gene	<809. .>4062	gene
gene	/gene="Pk2"	gene	/gene="Pk2"	gene
gene	/locus_tag="CG6504"	gene	/locus_tag="CG6504"	gene
gene	/map="3383-3383"	gene	/map="3383-3383"	gene
gene	/db_xref="FlYBASE:FBgn0041195"	gene	/db_xref="FlYBASE:FBgn0041195"	gene
gene	join(<809. .1568,1713. .2143,2212. .2603,2655. .2886,2950. .3177,3241. .3483,3542. .3789,3850. .3900,3954. .>4062	gene	join(<809. .1568,1713. .2143,2212. .2603,2655. .2886,2950. .3177,3241. .3483,3542. .3789,3850. .3900,3954. .>4062	gene
gene	/gene="Pk2"	gene	/gene="Pk2"	gene
gene	/locus_tag="CG6504"	gene	/locus_tag="CG6504"	gene
gene	/product="CG6504-RA"	gene	/product="CG6504-RA"	gene
gene	/db_xref="FlYBASE:FBgn0041195"	gene	/db_xref="FlYBASE:FBgn0041195"	gene
gene	join(809. .1568,1713. .2143,2212. .2603,2655. .2886,2950. .3177,3241. .3483,3542. .3789,3850. .3900,3954. .4062)	gene	join(809. .1568,1713. .2143,2212. .2603,2655. .2886,2950. .3177,3241. .3483,3542. .3789,3850. .3900,3954. .4062)	gene
gene	/gene="Pk2"	gene	/gene="Pk2"	gene
gene	/locus_tag="CG6504"	gene	/locus_tag="CG6504"	gene
gene	/codon_start=1	gene	/codon_start=1	gene
gene	/product="CG6504-PA"	gene	/product="CG6504-PA"	gene
gene	/protein_id="AAE53183.1"	gene	/protein_id="AAE53183.1"	gene
gene	/db_xref="GI:7297939"	gene	/db_xref="GI:7297939"	gene
gene	/db_xref="FlYBASE:FBgn0041195"	gene	/db_xref="FlYBASE:FBgn0041195"	gene
gene	/translation="MAQOGSGRSGSPGPPPPPPKPTTPGASPRGPTADSTPGDVTPRGPBPATKRVISGVGSTRAPAPADPDGPPPTAGKPVADGPGSTSPSLRDPGGSTDSPKSTSDKPSSTSDKPSSTPRTTAPTPDPPGSSSPRVSRTSGRTIFQEAAPS TSAKVVVLASBPAPKPEPEKGSFLPRKKKSGSSSTPSPSGHRAPSTIIAASALQNRKTNRLIYTTDEBVEALVFSVPFIILITSLVAPSVVGFPEKLTIVDDMDYLKYNKTLTGLDITFMDDEQANLTMSPPEGSQSGSGSGSGSGSPPEVGBAPPEAPEMDLQEDPKYNGNPLYDILRGHRAKRGSGVGDGNSGSDSPGNYLDDSDSIQANMS PHILBEGVPLIYENLLDGLRPRLQIRAKKESCYVNDATIRYNTCYAASGAEGRKPM HKSPFRTMHDLSTPIWTVALFRTGQYTNLDYDQRNKKIINDLKDTHLDRGSR LCVLEPFLFNENTDIPOSIGLIAEIPRTGVI PQAHQIVCMVSPFTRDSMLMTVIYI FVIMVIVYTIYEITTEIRKSGIKIYFCMSMILIDCALILGCVLALVNMISPFVMSL TAAHSAIVTQSGLDVLCFNMIIVYDMMALIAFLWMIKIFKISFNKTLVQPTTLKRC SWLAGSGIMFGIVFLAVYALQGLILPCKXPDPENFITSLITMTIMLIGDQVNLIEO ANRYLGLIYFLTYLILVFLFILNNFLAIIMETTYTKGELITQGSMSHGYIYRLKSCM LVYITLQGRRRHHPOASETEDKDAEDVGAADHETHEIRKNMTPAEQVYFKDIPQEG NOMVYLRNRRVGLHEILEIKLINMMDILKREKDYNNKKK"	gene	/translation="MAQOGSGRSGSPGPPPPPPKPTTPGASPRGPTADSTPGDVTPRGPBPATKRVISGVGSTRAPAPADPDGPPPTAGKPVADGPGSTSPSLRDPGGSTDSPKSTSDKPSSTSDKPSSTPRTTAPTPDPPGSSSPRVSRTSGRTIFQEAAPS TSAKVVVLASBPAPKPEPEKGSFLPRKKKSGSSSTPSPSGHRAPSTIIAASALQNRKTNRLIYTTDEBVEALVFSVPFIILITSLVAPSVVGFPEKLTIVDDMDYLKYNKTLTGLDITFMDDEQANLTMSPPEGSQSGSGSGSGSGSPPEVGBAPPEAPEMDLQEDPKYNGNPLYDILRGHRAKRGSGVGDGNSGSDSPGNYLDDSDSIQANMS PHILBEGVPLIYENLLDGLRPRLQIRAKKESCYVNDATIRYNTCYAASGAEGRKPM HKSPFRTMHDLSTPIWTVALFRTGQYTNLDYDQRNKKIINDLKDTHLDRGSR LCVLEPFLFNENTDIPOSIGLIAEIPRTGVI PQAHQIVCMVSPFTRDSMLMTVIYI FVIMVIVYTIYEITTEIRKSGIKIYFCMSMILIDCALILGCVLALVNMISPFVMSL TAAHSAIVTQSGLDVLCFNMIIVYDMMALIAFLWMIKIFKISFNKTLVQPTTLKRC SWLAGSGIMFGIVFLAVYALQGLILPCKXPDPENFITSLITMTIMLIGDQVNLIEO ANRYLGLIYFLTYLILVFLFILNNFLAIIMETTYTKGELITQGSMSHGYIYRLKSCM LVYITLQGRRRHHPOASETEDKDAEDVGAADHETHEIRKNMTPAEQVYFKDIPQEG NOMVYLRNRRVGLHEILEIKLINMMDILKREKDYNNKKK"	gene
gene	complement (5688. .25847)	gene	complement (5688. .25847)	gene
gene	/locus_tag="CG31764"	gene	/locus_tag="CG31764"	gene
gene	/map="3383-3384"	gene	/map="3383-3384"	gene
gene	/db_xref="FlYBASE:FBgn0051764"	gene	/db_xref="FlYBASE:FBgn0051764"	gene
gene	complement (join(5688. .6603,9611. .9992,9976. .10160,11882. .12049,24067. .24255,25749. .25847))	gene	complement (join(5688. .6603,9611. .9992,9976. .10160,11882. .12049,24067. .24255,25749. .25847))	gene
gene	/locus_tag="CG31764"	gene	/locus_tag="CG31764"	gene
gene	/product="CG31764-PA"	gene	/product="CG31764-PA"	gene
gene	/db_xref="FlYBASE:FBgn0051764"	gene	/db_xref="FlYBASE:FBgn0051764"	gene
gene	complement (join(6			

```

/codon_start=1
/product="CG31764-PA"
/protein_id="AAFS3185.2"
/db_xref="GI:22946332"
/db_xref="FLYBASE:FBgn0051764"
/transtable="MKLLFGLMTVIALSGFTLLPARNDDODDAVIVKPSRGPESD
FNRSGVIDTSGCYPELOPNPSSFNVGI.FDSFDLLRFLRLRPLVIGRSGSGPAP
TSDSDSDSDSGAGCFRVRPLIPLDSKNATTTSTIKVNHAKVINEITYGDSNVF
KRLNVNVPPLSGEERVAQGVHTSGDPLAPSPSTASPCAFEEFDEDDDDREPLLE
KOPQNEVRLIDEPKSTTPNTSYTNTLPHATPDISEAEADISIEDIVETQPEOKQT
MQLQOMMEPLLEEIRHEEDDEQHEDEYERKEERKEEFSSDDEGATTPVVM
SETFDNWEAFPODQDRHEDNEODDDHLDPEPNLNNEDIENFVPIDISNDIAYND
LAAAPNRPINDDAEFTVHPAIVRPMHREKLSLGEPAK"
complement(join(6007..6603,9611..9892,9976..10160,
11882..12049,24067..24130))

```

```

mRNA      complement(join(9415..9892,9976..10160,11882..12049,
24067..124267))
          /locus_tag="CG31764"
          /product="CG31764-RC"
          /db_xref="FLYBASE:Pgmn0051764"
          complement(join(9605..9892,9976..10160,11882..12049,
CDS

```

Query Match	16.7%;	Score 83.2;	DB 3;	Length 246230;
-------------	--------	-------------	-------	----------------

Matches 178; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

157 CCTCAATATACTATTGCTGCCCAATTATTCTGAGTCGGTGGATCTTATACTAC 216

46099 CCACTTGTGCTCATCTTCTCTCGCAGTTCGTCTGGGCGTGGCAACACCTTGTA

217 ACTTTAGGTGTATCCTATATGGACGATAACACCAAAATCTAAACACCGACTGATA 276

46039 TCCCTGGACAGACCTACCTCGACGACAACACCAAGAAGACGAACACTCCGCTCATGCTG 45988

277 AGTTTCTATTTCCTCGTATGCTCGGACCTGCAACCGGTATGCTTGGCTAGCGTC 336

b 45979 GCGTGCCATGGCACTCAGGATGATTGGACCAGTCGTGGGATCTTTTTCGGTTTCATA 4592

337 TGTCTAAGTTCACATTCGCCAACGTTGACGCCAACAAATTGATAATAATGATCCTAGA 396

b
45919 TCGCTTAACACCTTCATCGATCCAACCAAGACGCCGTTGATCGACAGCAAGACCCCGC 4586

397 TGGTANGACATGGTGGTTAAGTTGGGTGATACTAGGATCAACTTTAATATTTTTCGA 456

b
45859 TGGCTAGCTGCCTGGTGGCTCGGCTGGGTGATCCTGGGCACCCCTCATGTGCTCTTCTCC 45860

457 ACGTGGATTGTTTCTAAATTTCTGCCAAGAG 493

b 45799 GGTCTATCGGACTGTTCCCAAGCAATTGCCCAAG 45763

Q578543/c

DEFINITION Sequence 6301 from Patent WO0171042.

ERSION CQ578543.1 GI:41641008

SOURCE *Drosophila* sp.

ORGANISM	REFERENCE
Drosophila sp.	1
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
AUTHORS	Venter, J.C., Adams, M., Li, P.W., and Myers, E.W.
TITLE	Detection kits, such as nucleic acid arrays, for detecting the expression of 10,000 or more Drosophila genes and uses thereof
JOURNAL	Patent: WO 0171042-A 6301 27-SEP-2001; PE Corporation (NY) (US)
FEATURES	Location/Qualifiers
source	1..5349
	/organism="Drosophila sp." /mol_type="unassigned DNA" /db_xref="taxon:7242"

	Query Match	14.5%	Score 72.4	DB 6	Length 5349
	Best Local Similarity	58.0%	Pred. No. 2.3e-09		
	Matches 127	Conservative 0	Mismatches 92	Indels 0	Gaps 0
QY	276 AAGTTTTCTATATTTCTTGATGCTCGACCTGCAACCGGTTATGCTTGAGCTACCGT				335
DB	3451 AGGATGTCATCTTCTTAAGATGTGGGCCAGCATTTGTTACTTCCTAGCTGCTGT				3392
QY	336 CTGTCTTAAGTTCTACATTTCCGCAAGTCGACGCCAACAATTGATATATATGATTCGAG				395
DB	3391 CTGCTCGGCTTAATACATGAGCCACCTTGAGCCACTGATTTGGCGGAGAGATCCAG				3332
QY	396 ATGGTTANAGACATGATGGTTAAAGTTGGGTATCTAGATCAACATTATATATTTTTTGC				455
DB	3331 CTGGTTGGGGCCCTGGTGGCTGGGCTGGCTGCTGGCTGCATTTACCTTGGTCTCCG				3272
QY	456 AACGTGATTTGATTTGTTTCTATAAATTTCCCAAGAC				494
DB	3271 CATTCGCTGTTTATGTTCCCAAGACAGCTGCCAGTGC				3233

RESULT 15					
CQ578538					
LOCUS					
DEFINITION	CQ578538	2184 bp	DNA	linear	PAT 02-FEB-2004
ACCESSION	Sequence 6296 from Patent WO0171042.				
VERSION	CQ578538				
KEYWORDS	CQ578538.1 GI:41641005				
SOURCE	.				
ORGANISM	Drosophila sp.				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
AUTHORS	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
TITLE	Ephydroidea; Drosophilidae; Drosophila.				
JOURNAL	Venter,J.C., Adams,M., Li,P.W. and Myers,E.W. Detection kites, such as nucleic acid arrays, for detecting the expression of 10,000 or more Drosophila genes and uses thereof Patent: WO 0171042-A 6296 27-SEP-2001; PE Corporation (NY) (US)				
FEATURES	Location/Qualifiers				
source	1..2184 /organism="Drosophila sp." /mol_type="unassigned DNA" /db_xref="taxon:7242"				
ORIGIN					

Query Match	13.2%	Score 65.8	DB 6	Length 2184
Best Local Similarity	57.3%	Pred. NO. 1.7e-07		
Matches 118; Conservative	0	Mismatches 88	Indels 0	Gaps 0

Qy 289 TTTCTTGATGCTGGACCGTGCAACGGTATGCTCTGGCTAGCGTGTCTAAAGTTC 348
Db 889 TTTCTGGTATGCTGGGCTCTGCTATGGGCTTCTCCATGGATCTGTGTGCTCCGGCTG 948
Qy 349 TACATTCGCCAACGTTGAGGCCCAACAATTGATATATAATGATCTCTAAGATGGTTANGACA 408

Db 949 TATATGATCCCTTCAAAAAGCCACTGATTACCAAAAGATCCGCGCTGATGGCGCT 1008
 QY 409 TGGTGGTTAAGTTGGTGGTATCTAGATCACTTTAATATTTTGGCAAGTTGATGGA 468
 Db 1009 TGGTGGATAGTTGGATTTCTCTCACAATTCATCTGACTATCTCAGCAGTGTGCGGC 1068
 QY 469 TTGTTTCTTAAATTTGCCCAAGAGC 494
 Db 1069 ATGTTCCCAAAAGAAATGCCAAGGGC 1094

Search completed: February 26, 2005, 19:44:59
 Job time : 2373 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 17:05:10 ; Search time 511 Seconds
(without alignments)
5769.141 Million cell updates/sec

Title: US-10-621-901-26

Perfect score: 498

Sequence: 1 ggcctcggaagaagatgcctc.....aaattctgccaagagctgca 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: N_Geneseq_16Dec04:*
2: geneseqn1980s:*
3: geneseqn1990s:*
4: geneseqn2000s:*
5: geneseqn2001as:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	497	99.8	498	9 ACC72953	Acc72953 Cat flea
2	497	99.8	498	12 ADL09595	Adl09595 Cat flea
3	478.6	96.1	498	3 AAC93886	Aac93886 Cat flea
4	478.6	96.1	498	12 ADL09993	Adl09993 Cat flea
5	189.4	38.0	3015	4 ABL05971	Abi05971 Drosophila
6	125.4	25.2	6522	4 ABL05970	Abi05970 Drosophila
7	83.2	16.7	2486	4 ABL09403	Abi09403 Drosophila
8	83.2	16.7	6105	4 ABL09402	Abi09402 Drosophila
9	72.4	14.5	5349	4 ABL06040	Abi06040 Drosophila
10	65.8	13.2	2184	4 ABL06037	Abi06037 Drosophila
11	65.8	13.2	4399	4 ABL06036	Abi06036 Drosophila
12	59	11.8	2552	4 ABL06041	Abi06041 Drosophila
13	41.8	8.4	2172	10 ADC85132	Adc85132 Ion trans
14	41.8	8.4	2172	10 ADC85132	Adc85132 Ion trans
15	41.8	8.4	2175	12 ADM91341	Adm91341 Human tRNA
16	41.8	8.4	2251	10 ADC85155	Adc85155 Ion trans
17	41.8	8.4	2634	10 ADD37434	Add37434 Human tRNA
18	41.8	8.4	2634	13 ADR40142	Adr40142 Human tRNA
19	41.8	8.4	5054	13 ADR08230	Adr08230 Full length
20	40.2	8.1	8866	4 AAS45433	Aas45433 Chemical

21	40.2	8.1	8866	6 ABR28280	Abk28280 DNA trans
22	39.6	8.0	5882	6 ABL32545	Abi32545 Human imm
23	39.4	7.9	110000	3 AAF22305_03	Continuation (4 of
24	38.8	7.8	110000	2 AAT42063_14	Continuation (15 of
25	38.8	7.8	110000	2 AAT42063_15	Continuation (16 of
26	38.6	7.8	1737	10 ADC85178	Adc85178 Ion trans
27	38.6	7.8	2166	10 ADC85157	Adc85157 Ion trans
28	38.6	7.8	2169	10 ADC85181	Adc85181 Ion trans
29	38.4	7.7	15295	4 ABL7039	Abi7039 Human mus
30	38.4	7.7	15295	8 ABR60027	Abx60027 cDNA enco
31	38.4	7.7	15295	12 ADJ30777	Adj30777 Human mus
32	38.2	7.7	2000	8 ADA71938	Ada71938 Rice gene
33	37.6	7.6	3020	10 ADB54000	Adb54000 Human pro
34	37.4	7.5	443	6 ABL94091	Abi94091 Arabidops
35	37.4	7.5	3831	3 AAC48177	Aac48177 Arabidops
36	37.4	7.5	4746	2 AAV57455	Aav57455 Arabidops
37	37.4	7.5	4746	5 AAD03790	Ado3790 Arabidops
38	37.4	7.5	4746	10 ADD44387	Add44387 Arabidops
39	37.4	7.5	4747	2 AAT09019	Aat09019 Arabidops
40	37.4	7.5	4758	12 ADQ13602	Adq13602 Thalecres
41	37.4	7.5	6022	10 ADD44388	Add44388 Arabidops
42	37.4	7.5	6022	12 ADQ13603	Adq13603 Thalecres
43	37.4	7.5	6042	2 AAT09018	Aat09018 Arabidops
44	37.4	7.5	6172	2 AAV57454	Aav57454 Arabidops
45	37.4	7.5	6172	5 AAD03789	Ado3789 Arabidops

ALIGNMENTS

RESULT 1
ACCT2953 standard; DNA; 498 BP.
ACCT2953:
14-JUL-2003 (first entry)
Cat flea hindgut and Malpighian tube related gene #11.
Antiparasitic; gene therapy; vaccine; cat flea; head; nerve cord;
hindgut; malpighian tube; infestation; ss.
Ctenocephalides felis.
WO2003031577-A2.
17-APR-2003.
04-OCT-2002; 2002WO-US031878.
10-OCT-2001; 2001US-0328347P.
(HESK-) HESKA CORP.
Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;
WPI; 2003-393439/37.
New nucleic acid molecule, useful for preparing a composition for
treating or preventing flea infestation in a mammal.
Claim 1; Page 76; 247pp; English.
The invention relates to the isolation of nucleic acid sequence from cat
fleas that encode head, nerve cord, hindgut and malpighian tube proteins
or sequences that hybridize to these. The nucleic acids are useful for
preparing a composition for treating or preventing flea infestation in a
mammal. The proteins can also be used to raise antibodies for the
treatment of flea infestations in mammals. This sequence represents one
of the isolated sequences
Sequence 498 BP; 148 A; 102 C; 103 G; 144 T; 0 U; 1 Other;

Query Match	99.8%;	Score 497;	DB 9;	Length 498;
Best Local Similarity	100.0%;	Pred. NO. 3.7e-139;		
Matches 498; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	GGCCCTGGAACAAGATCCCTTGACATCACACACCGAAATGGAGCGTCTAGATCAAAAT	60
Db	1	GGCCCTGGAACAAGATGCCCTTGACATCACACCGAAATGGAGCGTCTAGATCAAAAT	60
QY	61	GCCACATTGAAAGTTCTTAATAAACAAGAGAGAAAATTGTGCGACATGATGCGAA	120
Db	61	GCCACATTGAAAGTTCTTAATAAACAAGAGAGAAAATTGTGCGACATGATGCGAA	120
QY	121	GGAGTTGAATGCGAAACAGAGAGGGAGCATGCGACCTCAATATATCTATTTTGCTGC	180
Db	121	GGAGTTGAATGCGAAACAGAGAGGGAGCATGCGACCTCAATATATCTATTTTGCTGC	180
QY	181	CAATTAAATTTCTGGAGTCGGTGGATCTTTATATCTACACTTTAGGTGATATCCATATAGAC	240
Db	181	CAATTAAATTTCTGGAGTCGGTGGATCTTTATATCTACACTTTAGGTGATATCCATATAGAC	240
QY	241	GATTAACACCAAAAAATCTTAAACACACAGACATGATAATTTTCTATTTCTTCGTATG	300
Db	241	GATTAACACCAAAAAATCTTAAACACACAGACATGATAATTTTCTATTTCTTCGTATG	300
QY	301	CTCGGACCTGCAACCGGTTATGCTTGGCTAGCGTCTGTCTTAAAGTTCTACATTTGCCA	360
Db	301	CTCGGACCTGCAACCGGTTATGCTTGGCTAGCGTCTGTCTTAAAGTTCTACATTTGCCA	360
QY	361	ACGTTAAGCCCAACAATTGATATATATATCTCTAATGGTTANAGCATGGTGGTTAAGT	420
Db	361	ACGTTAAGCCCAACAATTGATATATATATCTCTAATGGTTANAGCATGGTGGTTAAGT	420
QY	421	TGGGTGATCTAGAGATCAATTAAATATTTTGGCAAGTTGATGATGTTTCTTAA	480
Db	421	TGGGTGATCTAGAGATCAATTAAATATTTTGGCAAGTTGATGATGTTTCTTAA	480
QY	481	ATTCTGCCAAGACTGCA	498
Db	481	ATTCTGCCAAGACTGCA	498

XX	ADL09595	
XX	ADL09595 standard; cDNA; 498 BP.	
AC	ADL09595;	
XX		
DT	01-JUL-2004 (first entry)	
XX		
DE	Cat flea hindgut and malpighian tubule (HMT) protein cDNA #12.	
XX		
KW	flea; head and nerve cord protein; HNC;	
KM	hindgut and malpighian tubule protein; HMT; flea infestation;	
KM	anti-arthropod vaccine; chemotherapeutic drug; insecticide; gene; ss	
XX	cat flea.	
OS	Ctenocephalides felis.	
XX		
EN	US2004067516-A1.	
XX		
PD	08-APR-2004.	
XX		
PF	16-JUL-2003; 2003US-00621901.	
XX		
PR	22-JUL-2002; 2002US-0319414P.	
XX		
PA	(BRAN/) BRANDT K S.	
PA	(GAIN/) GAINES P J.	
PA	(STIN/) STINCComb D T.	
PA	(WISN/) WISNEWSKI N.	
XX		
TI	Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N; .	

XX WPI; 2004-304579/28 .
 DR
 XX Novel flea head and nerve cord protein and flea hindgut and malpighian
 PT tubule protein, useful for reducing flea infestations.
 PT
 XS Claim 5; SEQ ID NO 26; 35pp; English.
 XS

CC The invention relates to a flea head and nerve cord (HNC) protein and a
CC flea hindgut and malpighian tubule (HMT) protein. The invention also
CC relates to an isolated nucleic acid molecule expressed by a tissue chosen
CC from a flea HMT tissue and a flea HNC tissue, identified by a method
CC involving constructing a cDNA library enriched for HMT or HNC expressed
CC sequences and identifying a nucleic acid molecule in the library, and an
CC isolated antibody that selectively binds an HNC or HMT protein. The
CC proteins are useful for identifying compounds capable of inhibiting
CC activity of the proteins which involves contacting a protein with a
CC putative inhibitory compound under conditions in which, in the absence of
CC the compound, the protein has activity, and determining if the putative
CC inhibitory compound inhibits its activity. The proteins, nucleic acids
CC and antibodies are useful for reducing flea infestations. The proteins
CC and nucleic acids are useful as targets for anti-arthropod vaccines and
CC chemotherapeutic drugs. The proteins are useful for producing a
CC recombinant protein vaccine to protect an animal from flea infestation.
CC The antibodies are useful for passively immunising an animal in order to
CC protect the animal from fleas, as tools to screen expression libraries
CC and/or for recovering desired proteins from a mixture of proteins and
CC other contaminants. The antibodies are also useful for targeting
CC cytotoxic agents to fleas in order to directly kill such fleas. This
CC sequence represents cDNA encoding a cat flea HMT protein of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 498 BP; 148 A; 102 C; 103 G; 144 T; 0 U; 1 Other;

Query Match	99.8%	Score 497;	DB 12;	Length 498;
Best Local Similarity	100.0%;	Pred. No. 3.7e-139;		
Matches 498; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

Qy	1	GGCCCTGGACAGATGCCCTTCGACTCAACACCAATATGGAGCTGTCTACGATCAAAAT	60
Db	1	GGCCCTGGACAGATGCCCTTCGACTCAACACCAATATGGAGCTGTCTACGATCAAAAT	60
Qy	61	GCCACCATTTGAAGTTCTTTAATPAACAGAAAGGAAAACTTTGTGCCACTGATGGCGAA	120
Db	61	GCACACATTTGAAGTTCTTTAATPAACAGAAAGGAAAACTTTGTGCCACTGATGGCGAA	120
Qy	121	GGAGTTGAATGCGAAACAGAAAGGAGCGATCGCACCTCAATATATATATTGGCTGCC	180
Db	121	GGAGTTGAATGCGAAACAGAAAGGAGCGATCGCACCTCAATATATATATTGGCTGCC	180
Qy	181	CAATTATATTTCTGAGTCGGTGGATCTTTATATCAACTTTAGGTATCTATATGAC	240
Db	181	CAATTATATTTCTGAGTCGGTGGATCTTTATATCAACTTTAGGTATCTATATGAC	240
Qy	241	GATPAACCAAAAAATCTPAAACACCGACATGATAGTTTCTTATTTTCTTCGGTANG	300
Db	241	GATPAACCAAAAAATCTPAAACACCGACATGATAGTTTCTTATTTTCTTCGGTANG	300
Qy	301	CTCGACCTGCACACCGGTTATGCCCTTGGCTGACGCTGTCTPAAAGTTCTACATTTGCCA	360
Db	301	CTCGACCTGCACACCGGTTATGCCCTTGGCTGACGCTGTCTPAAAGTTCTACATTTGCCA	360
Qy	361	ACGTTGACGCCCACAATTGATPAATATGATTCCTAGATGGTTANGAGATGSGTAAAGT	420
Db	361	ACGTTGACGCCCACAATTGATPAATATGATTCCTAGATGGTTANGAGATGSGTAAAGT	420
Qy	421	TGGGTGATCTAGATCAACTTTATATTTTTCGACAGTTGATGGATTTGTTCTPAA	480
Db	421	TGGGTGATCTAGATCAACTTTATATTTTTCGACAGTTGATGGATTTGTTCTPAA	480
Qy	481	ATTCTGCCAAGAGCTGCA	498

Db 481 ATTCTGCCAAGAGCTGCA 498

RESULT 3
AAC93886
ID AAC93886 standard; cDNA; 498 BP.

XX AAC93886;

XX 19-FEB-2001 (first entry)

XX Cat flea hindgut and Malpighian tubule (HMT) cDNA, SEQ ID NO:381.

XX Cat flea, hindgut and Malpighian tubule nucleic acid; HMT;

XX flea infestation; vaccine; antiparasitic; therapeutic target; diagnosis;

XX detection; ss.

XX Ctenocephalides felis.

XX MO200061621-A2.

XX 19-OCT-2000.

XX 07-APR-2000; 2000WO-US009437.

XX 09-APR-1999; 99US-0128704P.

XX (HESK-) HESKA CORP.

XX Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;

XX WPI; 2000-656323/63.

XX flea Malpighian tubule and head and nerve cord tissue derived nucleic

XX acids useful for the prevention, diagnosis and treatment of flea

XX infestations.

XX Claim 26; Page 372; 964pp; English.

XX The invention relates to novel cat flea (Ctenocephalides felis) nucleic

XX acids which are expressed in hindgut and Malpighian tubule (HMT) tissue

XX or head and nerve cord (HNC) tissue. The invention also relates to the

XX encoded proteins. The invention additionally encompasses expression

XX constructs, recombinant viruses and recombinant cells comprising the

XX nucleic acids of the invention, recombinant production of the proteins,

XX antibodies against the proteins, a method of identifying inhibitors of

XX administration to an animal. The nucleic acids, and the proteins they

XX encode may be used in the prevention, treatment and diagnosis of diseases

XX associated with flea infestations. For example, the nucleic acids may be

XX used to produce an HMT or HNC protein according to standard recombinant

XX DNA methodology by inserting the nucleic acids into a host cell and

XX culturing the cell to express the protein. The HMT and HNC nucleic acids

XX may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect

XX and quantitate the presence of cat flea or other homologous nucleic acid

XX sequences in samples. They may also be used to study the expression and

XX function of the proteins and their role in metabolism. The HMT and HNC

XX proteins may be used as antigens in the production of specific

XX antibodies, and in assays to identify modulators (agonists and

XX antagonists) of HMT and/or HNC protein expression and activity. The anti-

XX HMT/HNC protein antibodies and antagonists may also be used to

XX downregulate protein expression and activity. The antibodies may also be

XX used as diagnostic agents for detecting the presence of flea polypeptides

XX in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The

XX present sequence represents a cat flea HMT cDNA of the invention

XX

XX Sequence 498 BP; 147 A; 101 C; 105 G; 145 T; 0 U; 0 Other;

XX Query Match 96.1%; Score 478.6; DB 3; Length 498;

XX Best Local Similarity 98.8%; Pred. No. 1.3e-133;

XX Matches 492; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 GGCCCTGAGACAGATGCCCTTGCATCCACCGAATATGAGCTGTCTACATCAAAAT 60
Db 1 GGCCCTGAGACAGATGCCCTTGCATCCACCGAATATGAGCTGTCTACATCAAAAT 60
QY 61 GCCACCATTTGAAGTTCTTAATTAACAGAGAGGAAAACTTTGGCCACACTGATGGGAA 120
Db 61 GCCACCATTTGAAGTTCTTAATTAACAGAGAGGAAAACTTTGGCCACACTGATGGGAA 120
QY 121 GGAGTTGAATGCGAAACAGAAAGGAGGATCGACCTCAATTAATCTATTGTGGCC 180
Db 121 GGAGTTGAATGCGAAACAGAAAGGAGGATCGACCTCAATTAATCTATTGTGGCC 180
QY 181 CAATTAATTTCTGAGTGGTGTGATCTTTATACATCACTTATAGTATCTATATGAC 240
Db 181 CAATTAATTTCTGAGTGGTGTGATCTTTATACATCACTTATAGTATCTATATGAC 240
QY 241 GATTAACCAAAAATCTTAATAACACGACCTGATTAAGTTTCTTATTTCTTCGATG 300
Db 241 GATTAACCAAAAATCTTAATAACACGACCTGATTAAGTTTCTTATTTCTTCGATG 300
QY 301 CTGGAGCTGCAACCGGTTATGCTGGCTAGGAGTGTCTAAGTCTACATTTGGCCA 360
Db 301 CTGGAGCTGCAACCGGTTATGCTGGCTAGGAGTGTCTAAGTCTACATTTGGCCA 360
QY 361 ACGTTGACGCCCAAAATTGATTAATGATCTAGATGTTANGAGCATGGTGAAGT 420
Db 361 ACGTTGACGCCCAAAATTGATTAATGATCTAGATGTTANGAGCATGGTGAAGT 420
QY 421 TGGGTGATTAAGTATCACTTAATATTTTGGCAAGTTGATTTGTTCTTAA 480
Db 421 TGGGTGATTAAGTATCACTTAATATTTTGGCAAGTTGATTTGTTCTTAA 480
QY 481 ATTCTGCCAAGAGCTGCA 498
Db 481 ATTCTGCCAAGAGCTGCA 497

RESULT 4

ID ADL09993 standard; cDNA; 498 BP.

XX ADL09993;

XX 01-JUL-2004 (first entry)

XX Cat flea hindgut and malpighian tubule (HMT) protein cDNA #410.

XX Flea; head and nerve cord protein; HNC;

XX hindgut and malpighian tubule protein; HMT; flea infestation;

XX anti-arthropod vaccine; chemotherapeutic drug; insecticide; gene; ss;

XX Ctenocephalides felis.

XX US2004067516-A1.

XX 08-APR-2004.

XX 16-JUL-2003; 2003US-00621901.

XX 22-JUL-2002; 2002US-0319414P.

XX (BRAN/) BRANDT K S.

XX (GAIN/) GAINES P J.

XX (STIN/) STINCHCOMB D T.

XX (WISN/) WISNIEWSKI N.

XX Brandt KS, Gaines PJ, Stinchcomb DT, Wisniewski N;

XX WPI; 2004-304579/28.

XX Novel flea head and nerve cord protein and flea hindgut and malpighian

XX tubule protein, useful for reducing flea infestations.

XX Claim 5; SEQ ID NO 426; 35pp; English.

CC The invention relates to a flea head and nerve cord (HNC) protein and a
CC flea hindgut and malpighian tubule (HMT) protein. The invention also
CC relates to an isolated nucleic acid molecule expressed by a tissue chosen
CC from a flea HMT tissue and a flea HNC tissue, identified by a method
CC involving constructing a cDNA library enriched for HMT or HNC expressed
CC sequences and identifying a nucleic acid molecule in the library, and an
CC isolated antibody that selectively binds an HNC or HMT protein. The
CC proteins are useful for identifying compounds capable of inhibiting
CC activity of the proteins which involves contacting a protein with a
CC putative inhibitory compound under conditions in which, in the absence of
CC the compound, the protein has activity, and determining if the putative
CC inhibitory compound inhibits its activity. The proteins, nucleic acids
CC and antibodies are useful for reducing flea infestations. The proteins
CC and nucleic acids are useful as targets for anti-arthropod vaccines and
CC chemotherapeutic drugs. The proteins are useful for producing a
CC recombinant protein vaccine to protect an animal from flea infestation.
CC The antibodies are useful for passively immunizing an animal in order to
CC protect the animal from fleas, as tools to screen expression libraries
CC and/or for recovering desired proteins from a mixture of proteins and
CC other contaminants. The antibodies are also useful for targeting
CC cytotoxic agents to fleas in order to directly kill such fleas. This
CC sequence represents cDNA encoding a cat flea HMT protein of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 498 BP; 147 A; 101 C; 105 G; 145 T; 0 U; 0 Other;

Query Match 96.1%; Score 478.6; DB 12; Length 498;
Best Local Similarity 98.8%; Pred. No. 1.3e-133;

Matches 492; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1 GGCCCTGGACAAGATGCGCTTCGACACGCAATATGAGTGTCTGACATCAAAAT 60
DB 1 GGCCCTGGACAAGATGCGCTTCGACACGCAATATGAGTGTCTGACATCAAAAT 60
QY 61 GCCACCAATGAACTTTTAATTAACAGAAAGAGAACTTTGTCACACATGATGCGAA 120
DB 61 GCCACCAATGAACTTTTAATTAACAGAAAGAGAACTTTGTCACACATGATGCGAA 120
QY 121 GGAAGTGAATGCGAAGACAGAAAGAGAGATGCGACCTCAATATATATATTTGCTGCC 180
DB 121 GGAAGTGAATGCGAAGACAGAAAGAGAGATGCGACCTCAATATATATATTTGCTGCC 180
QY 181 CAATTAATTTCTGAGTGGTGGATCTTTTATACACTTAGTGTATCTTAATATGAGAC 240
DB 181 CAATTAATTTCTGAGTGGTGGATCTTTTATACACTTAGTGTATCTTAATATGAGAC 240
QY 241 GATTAACACCAAAAATCTTAATAACACGACATGATGATTTTCTTATTTCTTCGATG 300
DB 241 GATTAACACCAAAAATCTTAATAACACGACATGATGATTTTCTTATTTCTTCGATG 300
QY 301 CTGGAACCTGCAACCGGTATGCTTGGCTAGCGCTGTCTGTAAGTTCTACATTTGCCA 360
DB 301 CTGGAACCTGCAACCGGTATGCTTGGCTAGCGCTGTCTGTAAGTTCTACATTTGCCA 360
QY 361 ACGTGAACGCAACAATGATTAATGATCTAGAGGTAGAGATGATGATGATGATGATGAT 420
DB 361 ACGTGAACGCAACAATGATTAATGATCTAGAGGTAGAGATGATGATGATGATGATGAT 420
QY 421 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
DB 421 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
QY 481 ATTCTGCAAGAGCTGCA 498
DB 481 A-TCTGCAAGAGCTGCA 497

RESULT 5

ABL05971
ID ABL05971 standard; cDNA; 3015 BP.

XX ABL05971;

AC 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 12395.

XX Drosophila; developmental biology; cell signaling; insecticide;

KW pharmaceutical; gene; ss.

XX Drosophila melanogaster.

OS WO200171042-A2.

XX 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2001; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EW;

PI WPI; 2001-656860/75.

DR P-P-SDB; ABB61866.

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.

PS Claim 1; SEQ ID NO 12395; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signaling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL01875) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 3015 BP; 754 A; 741 C; 799 G; 721 T; 0 U; 0 Other;

Query Match 38.0%; Score 189.4; DB 4; Length 3015;

Best Local Similarity 61.3%; Pred. No. 3.3e-46;

Matches 304; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

QY 1 GGCCCTGGACAAGATGCGCTTCGACACGCAATATGAGTGTCTGACATCAAAAT 60
DB 1127 GGAACCGGCGAGGATCAGCTGCTGACCTGGAATGCGCGGAGTCCCGATGAAAAAT 1186
QY 61 GCCACCAATGAACTTTTAATTAACAGAAAGAGAACTTTGTCACACATGATGCGAA 120
DB 1187 GCCACCAATGAACTTTTAATTAACAGAAAGAGATGCAACACCTATGCGCTTGAATGAGGA 1246
QY 121 GGAAGTGAATGCGAAGACAGAAAGAGAGATGCGACCTCAATATATATATTTGCTGCC 180
DB 1247 GGAAGTGAATGCGAAGACAGAAAGAGAGATGCGACCTCAATATATATATTTGCTGCC 180
QY 181 CAATTAATTTCTGAGTGGTGGATCTTTTATACACTTAGTGTATCTTAATATGAGAC 240
DB 1307 CAGTTCAATCTGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 1366
QY 241 GATTAACACCAAAAATCTTAATAACACGACATGATGATTTTCTTATTTCTTCGATG 300
DB 1367 GATTAATCTTAATAATCAAAATCTCGGCTTGTGATGCTTTCAATCTTTTGGCAGT 1426

Qy	301	CTCGGACCTGCACACCGGTTATGACCTTGGCTAGCGCTGCTCTAAAGTTCTTAACATTTCCGCA	360
Db	1427	TTGGGCTCTGCGCATTTGGATTGATGACCCCTGGCAATCTTTCTCTCGCGCTTAATATGACACCA	1486
Qy	361	ACGTTGACGCGCACAAATTTGATTAATTAATATCTAGATGATTANGACATGGTGTTAACT	420
Db	1487	CAAAATGATATCGGATGATCAACAACAAGATCCCGCTGCTGCGGCGCTGGTGGCTGGCG	1546
Qy	421	TGGGTGATTACTAGATCAACTTAATATTTTGTGACAGCTTGATTGATTGTTTCTTAA	480
Db	1547	TGGCTGGTGAATGGGGGCGCTGCTCTCTTCCTGGGAGTTTCTCTCCATGTTTCCCAA	1606
Qy	481	ATTCTGCCAAGAGCTG 496	
Db	1607	GAATTGCCAGAGCAG 1622	
RESULT 6			
ABL05970/c			
ID	ABL05970	standard; cDNA; 6522 BP.	
XX	XX	ABL05970;	
XX	AC		
XX	ABL05970;		
XX	26-MAR-2002	(first entry)	
XX	XX		
XX	Drosophila melanogaster	expressed polynucleotide SEQ ID NO 12392.	
XX	KM	Drosophila; developmental biology; cell signalling; insecticide;	
XX	KM	pharmaceutical; gene; ss.	
XX	OS	Drosophila melanogaster.	
XX	XX	MO200171042-A2.	
XX	PN		
XX	27-SEP-2001.		
XX	23-MAR-2001;	2001WO-US009231.	
XX	PR	23-MAR-2000; 2000US-0191637P.	
XX	PR	11-JUL-2000; 2000US-00614150.	
XX	PA	(PEKE) PE CORP NY.	
XX	PI	Venter JC, Adams M, Li PWD, Myers EW;	
XX	DR	WPI; 2001-656860/75.	
XX	DR	P-PSDB; ABB61867.	
XX	PT	New isolated nucleic acid detection reagent for detecting 1000 or more	
XX	PT	genes from Drosophila and for elucidating cell signalling and cell-cell	
XX	XX	interactions.	
XX	PS	Claim 1; SEQ ID NO 12392; 21pp + Sequence Listing; English.	
XX	XX		
XX	CC	The invention relates to an isolated nucleic acid detection reagent	
XX	CC	capable of detecting 1000 or more genes from Drosophila. The invention is	
XX	CC	useful in developmental biology and in elucidating cell signalling and	
XX	CC	cell-cell interactions in higher eukaryotes for the development of	
XX	CC	insecticides, therapeutics and pharmaceutical drugs. The invention	
XX	CC	discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA	
XX	CC	sequences (ABU01840-ABU16175) and the encoded proteins (ABBS7737-	
XX	CC	ABBS72072). The sequence data for this patent did not form part of the	
XX	CC	printed specification, but was obtained in electronic format directly	
XX	CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences	
XX	XX		
SO	Sequence 6522 BP; 1770 A; 1524 C; 1436 G; 1792 T; 0 U; 0 Other;		
Query Match			
25.2%; Score 125.4; DB 4; Length 6522;			
Best Local Similarity 65.6%; Pred. No. 8.3e-27;			
Matches 183; Conservative 0; Mismatches 96; Indels 0; Gaps 0;			
Qy	1	GGCCCTGGACAAGATGCCCTTGACCTACCAACCAATATGAGAGCTGTCTAGATTAAT	60

Db	3510	GGACCCGCGAGGATGACATTGGCTCTGACCTCGAGATTGGCGGGATGCCGATGAAAAT	3451
Qy	61	GGCACCAATTAAAGTTCTTTAATAACAGAAAGAAAACTTTGTGCCACACTGATGGGAA	120
Db	3450	GCCCAATATGAGGGCTATCGAGAGCACGATCCAAAGCCCTATCCGCTTGAAATGGAGA	3391
Qy	121	GGAGTTGAATGCGAAACGAAAGGAGCATCGACCTCAATATATCTATTTGCTGCC	180
Db	3390	GGAGCGGAGTCCGAGGTCGGTGAAGAAAATTGTGGCCCAACTTTTGCTCTTGATGCG	3311
Qy	181	CAATTAATTTTGGAGTCGGTGGATCTTTATACTACACTTTAGGTGATCTCTAATGAC	240
Db	3330	CAGTTCAATCTGAAATTTGAGGATCCCTATACACAGTTGGGAGTGTCTCAATGAC	3271
Qy	241	GATAACACCAAAAAATCTAAAAACCAACGACACTGATTAAGT	279
Db	3270	GATATATCTAATAAATCCAAAATCTCGGCTTTGTTGAGT	3232

Query Match	Beat Local Similarity	16.7%	Score 83.2	DB 4	Length 2486
Matches 169	Conservative 0	Mismatches 144	Indels 0	Gaps 0	

CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
CC sequences (AB101840-AB116175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/published_pct_sequences

XX Sequence 5349 BP; 1439 A; 1283 C; 1159 G; 1468 T; 0 U; 0 Other;

Query Match 14.5%; Score 72.4; DB 4; Length 5349;
Best Local Similarity 58.0%; Pred. No. 7e-11;
Matches 127; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 276 AAGTTTCTTATTTCTTGATGCTCGACCTGCAACCGGTTATGCTTGGCTACCGT 335

Db 3451 AGCAGTCTACCTCTTAAGATGCTGGGTCAGCCATTGATCTCCCTAGCTGCTG 3392

QY 336 CTGTCTAAATTTCTACCTTTCGCCAAGCTTGAGCCCAACATTTGATATATGCTTAG 395

Db 3391 CTGCTGCGGCTTACATTTAGCCACCTTGAGCCCACTGATGGGAGAGATCCACG 3332

QY 396 ATGATTAGACAGATGATGTTAGTGGGTGATCTAGATCACTTAATATTTTTCG 455

Db 3331 CTGATGGGGGCTGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 3272

QY 456 AACGTTGATTGATTTGTTCTTAAATTTGCGCAAGAC 494

Db 3271 CATTCTGCTGTTATGTTCCCAAGCAGTCCCAAGTGC 3233

RESULT 10
AB106037
ID ABL06037 standard; cDNA; 2184 BP.

XX ABL06037;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 12593.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX MO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PADB; ABB61934.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.

XX Claim 1; SEQ ID NO 12593; 21dp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
XX sequences (AB101840-AB116175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at http://wipo.int/pub/published_pct_sequences

XX Sequence 2184 BP; 516 A; 511 C; 557 G; 600 T; 0 U; 0 Other;

Query Match 13.2%; Score 65.8; DB 4; Length 2184;
Best Local Similarity 57.3%; Pred. No. 4.8e-09;
Matches 118; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 289 TTTCTTGATGCTGCGACCTGCAACCGGTTATGCTTGGCTAGCGCTCTTAAAGTTC 348

Db 889 TTTCTGATGCTGCGGCTGCTGATGCTTCCATGATCTTGGCTCCGCTG 948

QY 349 TACATTTGCGCAAGTTGACCGCAACATTTGATATATATATGATGATGTTAGACA 408

Db 949 TATATGATCTCTTCAAAAGCCATGATTAACCAACAGATCGGCTGGATGGCGCT 1008

QY 409 TGGTGTAAATTTGGGTGATCTAGATCACTTTAATTTTTCACAGTTGATGGA 468

Db 1009 TGGTGTAGATGTTGATTTCTTCACTTCATCTGATCTCAGCAGTGTCTGGGC 1068

QY 469 TTGTTCTTAAATTTGCTCCACAGC 494

Db 1069 ATGTTCCCAAGAAATGCCAAGGCG 1094

RESULT 11
AB106036/c
ID ABL06036 standard; cDNA; 4399 BP.

XX ABL06036;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 12590.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX MO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PADB; ABB61933.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.

XX Claim 1; SEQ ID NO 12590; 21dp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
XX sequences (AB101840-AB116175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 4399 BP; 1243 A; 1082 C; 953 G; 1121 T; 0 U; 0 Other;
 Query Match 13.2%; Score 65.8; DB 4; Length 4399;
 Best Local Similarity 57.3%; Pred. No. 6.3e-09;
 Matches 118; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
 QY 289 TTTCCTGATGCTCGGACCTGCAACGGTTATGCTTGGCTGACGCTCTCTAAAGTTC 348
 DB 2397 TTTCGATGATGCTGGGCTCTGCTATAGGCTTCTCCATGATGATCTTGTGCTCCGCTG 2338
 QY 349 TACATTTCCGCAACGTTGACGCCAACATGATATATATATCTAGATGTTANAGCA 408
 DB 2337 TATATGATTCCTCTTCAAAAAGCCATGATACCAAAACCATCCGGCTGATGGGCGCT 2278
 QY 409 TTGTTGGTAACTGGGATGACTAGATCAACTTTATATTTTGGCAAGTTGATGA 468
 DB 2277 TGGTGAATAGTTGATGATTTCTCTCACTTCATCTGATCTCAAGCATGTTGCTGGC 2218
 QY 469 TTGTTCTTAAATTCGCCAAGAGC 494
 DB 2217 ATGTTCCCAAGAAATGCCAAGGC 2192
 RESULT 12
 ABL06041
 ID ABL06041 standard; cDNA; 2552 BP.
 XX
 AC ABL06041;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 12605.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KM pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 XX
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW,
 XX
 DR WPI; 2001-656860/75.
 DR P-PsDB; ABB61938.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Claim 1; SEQ ID NO 12605; 21bp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB10511), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 CC

SO Sequence 2552 BP; 663 A; 625 C; 678 G; 586 T; 0 U; 0 Other;
 Query Match 11.8%; Score 59; DB 4; Length 2552;
 Best Local Similarity 67.5%; Pred. No. 5.7e-07;
 Matches 83; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
 QY 157 CCTCAATATATCTATTGTTGCTGCCAATTAATTTCTGAGTGGTGGATCTTATCTAC 216
 DB 609 CCCCAAGTGTCTCTGTTCTGCTCTCACTAATCTCGGAGTGGTCAAGCTCTGTTCTAT 668
 QY 217 ACTTAGGTGATACCTATATGACGATTAACCAAAAATCTTAAACACGACATGATA 276
 DB 669 ACCCTGGCATGACCTTACATGACGACATACCAAGCAAGCAACCTGCAATGCTG 728
 QY 277 AGT 279
 DB 729 ACT 731
 RESULT 13
 ADC85132
 ID ADC85132 standard; DNA; 2172 BP.
 XX
 AC ADC85132;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Ion transporter protein coding sequence, SEQ ID 2.
 XX
 DE Ion transporter protein coding sequence, SEQ ID 2.
 XX
 KW Ion transporter protein; kidney disease; liver disease; pancreas disease;
 KW immunological disease; thymus failure; reproductive disease;
 KW digestive disease; spleen disease; cancer; respiratory disease; myelitis;
 KW diabetes; hypertension; reperfusion injury; ischemia; retinitis;
 KW central nervous disease; skin disease;
 KW thyroid hormone-associated disease; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2172
 FT /*tag= a
 FT /product= "Ion transporter protein"
 XX
 PN WO2003074702-A1.
 XX
 PD 12-SEP-2003.
 XX
 PF 05-MAR-2003; 2003WO-JP002564.
 XX
 PR 06-MAR-2002; 2002JP-00061133.
 XX
 PR 01-APR-2002; 2002JP-00098852.
 XX
 PR 25-JUN-2002; 2002JP-00184883.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Nakanishi A, Hikichi Y, Uno Y;
 XX
 DR WPI; 2003-722073/68.
 DR P-PsDB; ADC85131.
 XX
 PT Novel marker proteins useful for the treatment and prevention of kidney
 PT disorders.
 XX
 PS Claim 10; SEQ ID NO 2; 186pp; Japanese.
 XX
 XX The present invention relates to novel ion transporter proteins
 CC (ADC85131, ADC85156, ADC85182 and ADC85184) and their coding sequences
 CC (ADC85131, ADC85157, ADC85183 and ADC85185). The proteins are useful as
 CC diagnostic markers for kidney diseases, liver diseases, pancreas
 CC diseases, immunological diseases associating thymus failures,
 CC reproductive diseases, digestive diseases, spleen diseases, cancer,
 CC respiratory diseases, myelitis, diabetes, hypertension, reperfusion
 CC injury following ischemia, retinitis, central nervous diseases, skin

CC diseases and thyroid hormone-associated diseases, and also for the
CC treatment and prevention of these diseases.
XX
SQ Sequence 2172 BP; 581 A; 451 C; 451 G; 689 T; 0 U; 0 Other;

Query Match 8.4%; Score 41.8; DB 10; Length 2172;
Best Local Similarity 47.8%; Pred. No. 0.08;
Matches 155; Conservative 0; Mismatches 163; Indels 6; Gaps 1;

QY 181 CAATTAATTTCTGGAGTCGGTGAATCTTATACACCTTAGTGATCTATATGAC 240
DB 706 CAATTAATTTCTGGAGTCGGTGAATCTTATACACCTTAGTGATCTATATGAC 240
QY 241 GATTAACACCAAAAATATCTAAACACCACTGATTAATTTCTTCTGATG 300
DB 766 GATTCTGTGCCCCACACACAGTCTTCTCTATATAGAACCGGTATGATGATC 825
QY 301 CTCGACCTGCAACCGGTATGCTTGGCTAGCGCTGTCTAAAGTCTTACATTCGCCA 360
DB 826 TTAGGCCCTGCTATGCTATGATGTATGGAGACCACTGATACATATACATGATGT 885
QY 361 ACGTTAGCCCAACAATTTGA-----TAAATAGATCTTATAGTGTANAGCATGTGG 414
DB 886 GCTATGGAGAAACACCTGATGATGATGATGATGATGATGATGATGATGATG 945
QY 415 TTAAGTTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 474
DB 946 ATTGGGTTCTTCTATCATGATGATGATGATGATGATGATGATGATGATGATG 1005
QY 475 CTTAAATTTCTGCCAAGAGCTGCA 498
DB 1006 CCAAAACATTTACCAAGTACAGCA 1029

RESULT 14
ADD37436
ID ADD37436 standard; cDNA; 2172 BP.

AC ADD37436;

DT 15-JAN-2004 (first entry)

DE Human transporter OAT-5 cDNA #2.

XX

KW Human; 86; gene; transporter; cytosolic; anorectic; antidiabetic;

KW anticonvulsant; gene therapy; PGC-1 associated disorder; liver tumour;

XX obesity; epilepsy; diabetes.

OS Homo sapiens.

PN US2003143675-A1.

PD 31-JUL-2003.

PF 22-MAY-2002; 2002US-00154419.

PR 12-MAY-2000; 2000US-0204211P.

PR 29-JUN-2000; 2000US-0215376P.

PR 31-JUL-2000; 2000US-0221769P.

PR 19-SEP-2000; 2000US-0233790P.

PR 25-SEP-2000; 2000US-0235107P.

PR 05-OCT-2000; 2000US-0238336P.

PR 14-NOV-2000; 2000US-0248354P.

PR 15-NOV-2000; 2000US-0248787P.

PR 15-DEC-2000; 2000US-0256240P.

PR 21-DEC-2000; 2000US-0258028P.

PR 22-JAN-2001; 2001US-0263169P.

PR 14-MAY-2001; 2001US-00858194.

PR 29-JUN-2001; 2001US-00895811.

PR 31-JUL-2001; 2001US-00919781.

PR 19-SEP-2001; 2001US-00957664.

PR 25-SEP-2001; 2001US-00964295.

PR 05-OCT-2001; 2001US-00972724.
PR 14-NOV-2001; 2001US-00002769.
PR 17-DEC-2001; 2001US-00024623.
PR 22-JAN-2002; 2002US-00055025.

XX (MILL-) MILLENNIUM PHARM INC.

PA Curtis RAJ, Gluckemann MA, Meyers RE;

XX WPI: 2003-851783/79.

DR P-PSDB; ADD37435.

XX New isolated nucleic acid, useful for preparing a composition for
PT treating PGC-1 associated disorders e.g. liver tumors, obesity, epilepsy
or diabetes.

PS Claim 1; SEQ ID NO 9; 663bp; English.

XX The invention relates to an isolated nucleic acid comprising a cDNA
CC encoding a human transporter protein, or its complement, a sequence that
CC is 60 % identical to the cDNA, a fragment comprising at least 30
CC nucleotides of the cDNA, or a sequence encoding a fragment of the
CC polypeptide comprising at least 10 contiguous amino acid residues of the
CC cDNA. Also included are a vector comprising the novel nucleic acid
CC molecule, producing the polypeptide, the isolated transporter
CC polypeptide, an isolated antibody that specifically binds to the
CC polypeptide, detecting the presence of the polypeptide or nucleic acid in
CC a sample, a kit, identifying a compound that binds to, or that modulates
CC the activity of, the polypeptide, and modulating the activity of the
CC polypeptide. The nucleic acid is useful for preparing a composition for
CC treating PGC-1 (not defined) associated disorders e.g. liver tumors,
CC obesity, epilepsy or diabetes. The present sequence encodes a novel human
CC transporter protein.

XX Sequence 2172 BP; 580 A; 450 C; 453 G; 689 T; 0 U; 0 Other;

Query Match 8.4%; Score 41.8; DB 10; Length 2172;
Best Local Similarity 47.8%; Pred. No. 0.08;
Matches 155; Conservative 0; Mismatches 163; Indels 6; Gaps 1;

QY 181 CAATTAATTTCTGGAGTCGGTGAATCTTATACACCTTAGTGATCTATATGAC 240

DB 706 CAATTAATTTCTGGAGTCGGTGAATCTTATACACCTTAGTGATCTATATGAC 240

QY 241 GATTAACACCAAAAATATCTAAACACCACTGATTAATTTCTTCTGATG 300

DB 766 GATTCTGTGCCCCACACACAGTCTTCTCTATATAGAAACCGGTATGATGATC 825

QY 301 CTCGACCTGCAACCGGTATGCTTGGCTAGCGCTGTCTAAAGTCTTACATTCGCCA 360

DB 826 TTAGGCCCTGCTATGCTATGATGTATGGAGACCACTGATACATATACATGATGT 885

QY 361 ACGTTAGCCCAACAATTTGA-----TAAATAGATCTTATAGTGTANAGCATGTGG 414

DB 886 GCTATGGAGAAACACCTGATGATGATGATGATGATGATGATGATGATGATG 945

QY 415 TTAAGTTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 474

DB 946 ATTGGGTTCTTCTATCATGATGATGATGATGATGATGATGATGATGATGATG 1005

QY 475 CTTAAATTTCTGCCAAGAGCTGCA 498

DB 1006 CCAAAACATTTACCAAGTACAGCA 1029

RESULT 15
ADM91341
ID ADM91341 standard; DNA; 2175 BP.

AC ADM91341;

DT 03-JUN-2004 (first entry)

DE Human organic ion transporter protein OATP-M1 gene sequence SegID3.
 XX organic anion transporter; kidney cell; medicine excretion;
 KM medicine re-absorption; medicine uptake; gene; ds; human; OATP-M1.
 XX
 OS Homo sapiens.
 Key Location/Qualifiers
 CDS 1..2175
 FT /*tag= a
 FT /product= "Human organic ion transporter (OATP-M1)"
 XX
 PN JP2004065086-A.
 XX
 PD 04-MAR-2004.
 XX
 PF 06-AUG-2002; 2002JP-00228131.
 XX
 PR 06-AUG-2002; 2002JP-00228131.
 XX
 PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 XX
 DR WPI: 2004-209112/20.
 DR P-PSDB; ADM91339.
 XX
 PT Novel organic anion transporter concerned in excretion and re-absorption
 PT of medicine in kidney, expressing in kidney, useful for measuring uptake
 PT and excretion of target substance in cell.
 XX
 PS Claim 7; SEQ ID NO 3; 29pp; Japanese.
 XX
 CC This invention relates to a novel organic anion transporter (OATP-M1)
 CC expressed in kidney cells and involved in excretion and re-absorption of
 CC medicine in the kidney. The invention may be useful for measuring the
 CC uptake and excretion of target substance in a cell, and for screening a
 CC substance which can uptake large quantities in a cell. The invention
 CC enables to measure the uptake and excretion of a target substance in the
 CC kidney specifically. The present sequence is that of a gene which encodes
 CC a human organic anion transporter protein of the invention.
 XX
 SQ Sequence 2175 BP; 583 A; 450 C; 452 G; 690 T; 0 U; 0 Other;

Query Match 8.4%; Score 41.8; DB 12; Length 2175;
 Best Local Similarity 47.8%; Pred. No. 0.08;
 Matches 155; Conservative 0; Mismatches 163; Indels 6; Gaps 1;

QY 181 CAATTAATTTCTGAGTCGGTGGATCTTTATACACCTTTAGGTATCTCTATATGAC 240
 DB |||||
 DB 706 CAACTATGTCTGGGGGAGGAGAACTCTCTTTATCTCTGGAGACGCTTCTTGAT 765
 QY 241 GATAACACCAAAAATCTAAACACAGCACTGATTAAGTTTCTTATTTCTTCTGATG 300
 DB |||||
 DB 766 GATTCTGTGCCCACACAAAGTCTTCTCTATATAGGAACGGTTATGCTATGCAATC 825
 QY 301 CTGGACCTGCAACCGGTTATGCTTGGCTAGCGCTGTCTTAAAGTTCTACATTTGCCA 360
 DB |||||
 DB 826 TTAGGCCCTGCTATTTGGCTATGTATGGAGACAACGCTAACATATACATGATGTT 885
 QY 361 ACGTTACGCCACAAATTTGA-----TAAATATGATCTTATGATGTTANAGACATGTTG 414
 DB |||||
 DB 886 GCTATGAGAGAAAGCACTGATGTCACATGAGATGATCCGATGTTGGAGCTTGGTGG 945
 QY 415 TTAAGTTGGGTGATACATGATGATCACTTAAATATTTTGAACGTTGATGATGTTT 474
 DB |||||
 DB 946 ATTGGGTTTCTTCTATATGATGATCTTTGCTTGTGCTTTAATATATACCTTTTCTGCTT 1005
 QY 475 CTTAAATTTCTGCCAGAGCTGCA 498
 DB |||||
 DB 1006 CCAAAACATTTACAGGTACAGCA 1029

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 18:50:01 ; Search time 147 Seconds

(without alignments)
5543.305 Million cell updates/sec

Title: US-10-621-901-26

Perfect score: 498

Sequence: 1 ggcctcggaacagatgcctc.....aaattctgcaagagctgca 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodaca/1/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodaca/1/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodaca/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodaca/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodaca/1/ina/PTUS COMB.seq:*
- 6: /cgn2_6/ptodaca/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	38.8	7.8	1830121	4 US-09-557-884-1	Sequence 1, Appl1
C 2	38.8	7.8	1830121	4 US-09-643-990A-1	Sequence 1, Appl1
C 3	38.4	7.7	601	4 US-09-949-016-126480	Sequence 126480, A
C 4	38.4	7.7	69642	4 US-09-949-016-15339	Sequence 15339, A
C 5	37.6	7.6	2721	4 US-09-949-016-713	Sequence 713, App
C 6	37.6	7.6	2721	4 US-09-949-016-2163	Sequence 2163, App
C 7	37.6	7.6	2721	4 US-09-949-016-2164	Sequence 2164, App
C 8	37.6	7.6	2741	4 US-09-949-016-1119	Sequence 1119, App
C 9	37.4	7.5	4746	2 US-08-819-288-2	Sequence 2, Appl1
C 10	37.4	7.5	4746	2 US-08-400-348-2	Sequence 2, Appl1
C 11	37.4	7.5	4747	1 US-08-261-822A-2	Sequence 2, Appl1
C 12	37.4	7.5	4747	5 PCT-US95-07744A-2	Sequence 2, Appl1
C 13	37.4	7.5	6042	1 US-08-261-822A-1	Sequence 1, Appl1
C 14	37.4	7.5	6042	5 PCT-US95-07744A-1	Sequence 1, Appl1
C 15	37.4	7.5	6172	2 US-08-819-288-1	Sequence 1, Appl1
C 16	37.4	7.5	6172	3 US-08-400-348-1	Sequence 1, Appl1
C 17	35.4	7.1	268223	4 US-09-596-002-41	Sequence 41, Appl1
C 18	34.6	6.9	168575	3 US-09-426-280-1	Sequence 1, Appl1
C 19	34.2	6.9	601	4 US-09-949-016-195940	Sequence 195940, A
C 20	34.2	6.9	601	4 US-09-949-016-195941	Sequence 195941, A
C 21	34.2	6.9	3692	4 US-09-575-081B-7	Sequence 7, Appl1
C 22	34.2	6.9	3725	4 US-09-799-451-260	Sequence 260, App
C 23	34.2	6.9	231672	4 US-09-949-016-17296	Sequence 17296, A
C 24	34.2	6.9	251682	4 US-09-949-016-11973	Sequence 11973, A
C 25	34.2	6.9	325791	4 US-09-768-185A-1	Sequence 1, Appl1
C 26	34	6.8	150780	4 US-09-949-016-14711	Sequence 14711, A
C 27	33.8	6.8	928	4 US-09-270-767-1100	Sequence 1100, App

C 28	33.8	6.8	928	4 US-09-270-767-16382	Sequence 16382, A
C 29	32.8	6.6	601	4 US-09-949-016-134819	Sequence 134819, App
C 30	32.8	6.6	832	4 US-09-621-976-2813	Sequence 2813, App
C 31	32.8	6.6	903	4 US-09-328-352-3071	Sequence 3071, App
C 32	32.8	6.6	3183	4 US-09-248-796A-3959	Sequence 3959, App
C 33	32.8	6.6	76610	4 US-09-949-016-15521	Sequence 15521, App
C 34	32.8	6.6	125192	4 US-09-949-016-14120	Sequence 14120, App
C 35	32.8	6.6	421491	4 US-09-949-016-12805	Sequence 12805, App
C 36	32.8	6.6	421494	4 US-09-949-016-14060	Sequence 14060, App
C 37	32.8	6.6	678533	4 US-09-949-016-14577	Sequence 14577, App
C 38	32.8	6.6	678533	4 US-09-949-016-14578	Sequence 14578, App
C 39	32.6	6.5	474	4 US-09-621-976-18033	Sequence 18033, App
C 40	32.6	6.5	601	4 US-09-949-016-78977	Sequence 78977, App
C 41	32.6	6.5	601	4 US-09-949-016-78978	Sequence 78978, App
C 42	32.6	6.5	2108	4 US-09-270-767-12398	Sequence 12398, App
C 43	32.6	6.5	132871	4 US-09-949-016-13863	Sequence 13863, App
C 44	32.6	6.5	151607	4 US-09-949-016-12210	Sequence 12210, App
C 45	32.6	6.5	784019	4 US-09-949-016-14033	Sequence 14033, App

ALIGNMENTS

RESULT 1
US-09-557-884-1/c
Sequence 1, Application US/0957884
Patent No. 6506581
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/557,884
FILING DATE: 25-Apr-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: JUN-5-1995
ATTORNEY/AGENT INFORMATION:
NAME: Michelle S. Marks
REGISTRATION NUMBER: 41,971
REFERENCE/DOCKET NUMBER: P186P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
TELEFAX: 301-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1
Query Match 7.8%; Score 38.8; DB 4; Length 1830121;
Best Local Similarity 49.3%; Pred. No. 2.2; Indels 0; Gaps 0;
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
Oy 262 ACACGACGATGAAGTTTCTTATTTCTTCTGATGCTCGAGCTGCAACGCTTAT 321

Db 1506177 AAAGTGAAGTTTAAAGCCTTTGATTATGCGTGAAGAAATTACGCAAGAGGTTAT 1506118
Qy 322 GCCTTGCTAGCGCTCTGTCTAAAGTTCTACATTTGCGCAACGTTGACGCAACATTTAT 381
Db 1506117 AGCTTATGATATAATATAAGCATTTGCTGATGACATTAAGCAAGAAAGCATTTGCG 1506058
Qy 382 AATAATGATCTTATGATGTTTANGACATGCTGTTAAGTTGGTGATCTAGATCAACT 441
Db 1506057 AGTTTATGTGTAAAGGCTTGAGGAGATGGAAGGCGCATTAAGCCTGTGTAAGCAACT 1505998
Qy 442 TTAATATTTTTCGCAAGTTGAT 464
Db 1505997 TTAAACTTTTGAAGATAGAT 1505975

RESULT 2
US-09-643-990A-1/C
Sequence 1, Application US/09643990A
Patent No. 6528289
GENERAL INFORMATION:
APPLICANT: Robert D. Fleischmann
Mark D. Adams
Owen White
Hamilton O. Smith
J. Craig Venter
TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville,
STATE: MD
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2 inch diskette
COMPUTER: Dell Pentium
OPERATING SYSTEM: MS DOS V6.22
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/643,990A
FILING DATE: 23-Aug-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: 1995-06-07
APPLICATION NUMBER: 08/426,787
FILING DATE: 1995-04-21
ATTORNEY/AGENT INFORMATION:
NAME: Kenley K. Hoover
REGISTRATION NUMBER: 40,302
REFERENCE/DOCKET NUMBER: PB186P1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-610-5790
TELEFAX: 310-309-8439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1830121 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match 7.8%; Score 38.8; DB 4; Length 1830121;
Best Local Similarity 49.3%; Pred. No. 2.2;
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;
Qy 262 ACACCGACATGATAGTTTCTTATTTCTTCTGCTGCGAGCTGCAACCGGTTAT 321

Db 1506177 AAAGTGAAGTTTAAAGCCTTTGATTATGCGTGAAGAAATTACAGCAAGAGGTTAT 1506118
Qy 322 GCCTTGCTAGCGCTCTGTCTAAAGTTCTACATTTGCGCAACGTTGACGCAACATTTAT 381
Db 1506117 AGCTTATGATATAATATAAGCATTTGCTGATGACATTAAGCAAGAAAGCATTTGCG 1506058
Qy 382 AATAATGATCTTATGATGTTTANGACATGCTGTTAAGTTGGTGATCTAGATCAACT 441
Db 1506057 AGTTTATGTGTAAAGGCTTGAGGAGATGGAAGGCGCATTAAGCCTGTGTAAGCAACT 1505998
Qy 442 TTAATATTTTTCGCAAGTTGAT 464
Db 1505997 TTAAACTTTTGAAGATAGAT 1505975

RESULT 3
US-09-949-016-126480
Sequence 126480, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 126480
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-126480

Query Match 7.7%; Score 38.4; DB 4; Length 601;
Best Local Similarity 60.0%; Pred. No. 0.047;
Matches 63; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 374 CAATTGATATAATATCTCTAGATGTTTANGACATGCTGTTAAGTTGGTGATCTAG 433
Db 422 CAATTGATTAAGAAATAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
Qy 434 GATCAACTTAAATATTTTTCGACAGCTGATGATGATGATGATGATGATGATGATGATGAT 478
Db 482 GAACAAGTTAACTACTTTTAAAGCCTTTATTAAGATGATGATGATGATGATGATGATGATGAT 526

RESULT 4
US-09-949-016-15339
Sequence 15339, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15339

```

; LENGTH: 69642
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15339

```

Query Match	7.74;	Score 38.4;	DB 4;	Length 69642;
Best Local Similarity	60.04;	Pred. No. 0.55;		
Matches	63;	Conservative	0;	Mismatches 42; Indels 0;

374 CAATTGATTAATATGATCCTAGACGTGATGAGCATCTAG 433
Oy
35169 CAATTGATTGAAGATCTGCTGAGCATGCTTAATCTTTGATAGAAATTAATTA 352228
Db

QY 434 GATCACTTAAATTTTTCGACGGTGATTGGATGTTCCTA 478
 ||| ||| | |||| | ||| ||| |||
Db 35229 GAACAAGTTAACTACTTTTAAAGCCTTATTAGATGTTATCTA 35273

RESULT 5
US-09-949-016-713
; Sequence 713, Application US/09949016

```

: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: WITH HUMAN IMMUNE DEFICIENCY VIRUS

```

```

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/261 755

```

; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 713

```

```

; LENGTH: 2/21
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-713

```

Query Match	7.6%	Score 37.6;	DB 4;	Length 2721;
Best Local Similarity	47.6%;	Pred. No. 0.19;		
Matches 152; Conservative	0;	Mismatches 155;	Indels 12;	Gaps 1

Oy 192 TGGAGTCGGTGGATCTTATACACCTTAGTGATCTATATGAGCAGTAAACACCA 251
 Db 557 TGGATGGGGTAACTCCCATCTCTGCTTGGGATTTCTATATAGAGAGTTTGGCCAA 616

QY 252 AAAATCTAAACACCGAGCTGATAGTTTCTTATTTCTTCGTAGCTGGACCTGC 311

Db 617 ATTGAAATTTCTCTTTATATATTTGGGCTGTAGAAACAGAGCATTTATTTGCTCTT 676

QY 312 AACCGTTATGCCCTGGCTGACGGTGTCTAAAGTTCTACATTTGCCCAAGTTGACGC 371

Db 677 GATTGGACATTTTGTGGCATCTTCTGTGCAAAATGTTTAATGTGACACTCGATTTGTGAA 736

Qy	372	AAACAATTGATAATA-----TGATCCTGATGAGTTANAGCATGCTGTTAG	419
Db	737	CACGATGATCTGATCATAACTCCCACTGACACTGTTGGTCGGTGCATGGTGGTTGG	796

Oy 420 TTGGGTGATCTAGCATCTTATATTTTGGCAACGTGATGTGTTTCTTAA 479
 797 CTTCCTGATTTTGGCAGAGTTACGTGCTCATCTGCATTCCTTTTCTTTTGGCCAA 856
 Db

Oy 480 AATCTGCCAAGAGCTGCA 498
| | | | | | | |
Db 857 CACACTTCCAAAGGAAGA 875

RESULT 6

US-09-949-016-2163
Sequence 2163, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

;
; CURRENT APPLICATION NUMBER: US/09/949,016
;
; CURRENT FILING DATE: 2000-04-14
;
; PRIOR APPLICATION NUMBER: 60/241,755
;
; PRIOR FILING DATE: 2000-10-20

;; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2163
; LENGTH: 2721
```

TYPE: DNA
ORGANISM: Human
US-09-949-016-2163

Query match	7.08;	DB 4;	Length 222;
Best Local Similarity	47.6%;	Score No. 0.19;	
Matches 152; Conservative	0;	Mismatches 155;	Indels 12; Gaps 1.

DB 192 TGGAGTGGGTGAAACTCCCATCCCTGCTTGGGTATTTCTATATAGAAGATTTTGC 616

Db 617 ATTGAAATTCCTTATATATTTGGCTGTAGAAACAGAGACTATTATTGCTCTTT 676

Db 677 GATTGGACCTTTGTTGGCAATCATTCCTGTGCAAAAGTTTAAGTTGACACACGAGATTGTGAA 736

737 CACAGATGATCTGATCATTAAGTCCCACTGACATCTGTTGGTCCGTGCATGGTGGTTGG 796

Db 797 CTTCTGATTTGTCAGAGTTAAGCTGCTACCTGCGCCATTCTTTTCTTTTGGCCCA 856

Db 857 CACACTTCCAAGGAGGA 875

RESULT 7
US-09-949-016-2164
; Sequence 2164, Application US/09949016

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE METHODS OF DETECTION AND USES THEREOF

```

```

; FILE REFERENCE: CL001307
;
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
;
; PRIOR APPLICATION NUMBER: 60/241,755

```

;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498

```

; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2164

```

; SEQ ID NO 2164

LENGTH: 2721
TYPE: DNA
ORGANISM: Human
US-09-949-016-2164

Query Match
Best Local Similarity 47.6%; Pred. No. 0.19; Length 2721;
Matches 152; Conservative 0; Mismatches 155; Indels 12; Gaps 1;

192 TGGAGTCGGTGAATCTTTATATACACTTTAGGTGATTCCTATATGAGACATAACCA 251
|||
557 TGGAAATGGGGGAAACTCCATCCGCTTTGGGATTTCTATATAGAAATTTTGGCAA 616
|||
252 AAAATCTAAACACGACACTGATTAAGTTTCTTATTTCTTGATTCGTGACCTGC 311
|||
617 ATTTGAAATTCCTTTATATATATGCGCTTGAACAGAGACTATTTGGTCCCTT 676
|||
312 AACCGGTATGCTGGCTGAGCTGCTAAGTTCTAATTCGCAACGTGAGCC 371
|||
677 GATTTGACCTTTGCTGATCATCTCTGCAAAAGTTATGACACTGGAATTTGAA 736
|||
372 AACAAATGATATATA-----TGATCTAGATGCTTANAGACATGCTTAA 419
|||
737 CACAGATGATCTGATCATTAATCCCACTGACACTGCTGGGTGGTGCATGGTGG 796
|||
420 TTGGGTGATACAGATCACTTAATATTTTTCACACCTGATGATTTCTCTAA 479
|||
797 CTTTCTGATTTGTGACAGAGTTAACGTGCTACCTGCAATCTTTTCTTTGCCCA 856
|||
480 AATTCGCCAAGAGCTGCA 498
|||
857 CACACTTCCAAAGAGAGA 875
|||

RESULT 8

US-09-949-016-1119
Sequence 1119, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1119
LENGTH: 2741
TYPE: DNA
ORGANISM: Human
US-09-949-016-1119

Query Match
Best Local Similarity 7.6%; Score 37.6; DB 4; Length 2741;
Matches 152; Conservative 0; Mismatches 155; Indels 12; Gaps 1;

192 TGGAGTCGGTGAATCTTTATATACACTTTAGGTGATTCCTATATGAGACATAACCA 251
|||
577 TGGAAATGGGGGAAACTCCATCCGCTTTGGGATTTCTATATAGAAATTTTGGCAA 636
|||
252 AAAATCTAAACACGACACTGATTAAGTTTCTTATTTCTTGATTCGTGACCTGC 311
|||
637 ATTTGAAATTCCTTTATATATATGCGCTTGAACAGAGACTATTTATGGTCCCTT 696
|||
312 AACCGGTATGCTGGCTGAGCTGCTAAGTTCTAATTCGCAACGTGAGCC 371
|||

697 GATTGACCTTTGTGTCATCATTCCTGCAAAATGTTTATGTTGACACTGATTTGGAA 756
|||
372 AACATTTGATATATA-----TGATCTAGATGCTTANAGACATGCTGTTAA 419
|||
757 CACAGATGATCTGATCATTAATCCCACTGACACTGCTGGGTGGTGCATGGTGGTGG 816
|||
420 TTGGGTGATACAGATCACTTAATATTTTTCGACAGTTGATGATGTTCTCTAA 479
|||
817 CTTTCTGATTTGTGACGAGTTAACGTGCTCACTGCAATTCCTTTTCTTTGCCCA 876
|||
480 AATTCGCCAAGAGCTGCA 498
|||
877 CACACTTCCAAAGAGAGA 895
|||

RESULT 9

US-08-819-288-2
Sequence 2, Application US/08819288
Patent No. 595652
GENERAL INFORMATION:
APPLICANT: Ecker, Joseph
TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE
TITLE OF INVENTION: AND PATHOGENS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 595652zris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/819,288
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: UPN-2949
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4746 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 584..4468
US-08-819-288-2

Query Match
Best Local Similarity 7.5%; Score 37.4; DB 2; Length 4746;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

74 TTCTTAATAAACAAGAGAAACTTTGTGCCACTGATGGCGAAGATTGAATCG 133
|||
2391 TTCTTGACATGGGAATAACAGCAAGTTATGAAAGAGTGTGAAGGGGTTCAATGG 2450
|||
134 AACAGAGAAGGAGGACATGCGACCTCAATATATACATTTTGTGCGCAAT 184
|||
2451 AACAGAGAAGGAGGACATGCGACCTCAATATATACATTTTGTGCGCAAT 2501
|||

FILING DATE: June 17, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4747 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US95-07744A-2

Query Match 7.5%; Score 37.4; DB 5; Length 4747;
Best Local Similarity 58.6%; Pred. No. 0.29;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 74 TTCTTATATAACAGAGAGAAAACCTTGCCACACTGATGGGAGAGATTGAATGCG 133
Db 2391 TTGTTACATGAGAAATTAACAGCAAGTTTATTGAAAAGAGATGTTGAAGGGGTTTCATGGG 2450
Qy 134 AAACAGAGAAGGAGAGATGCACCTCAATAATATCTATTGCTGCCCAAT 184
Db 2451 AAACAGAGAAGGAGATGCACCAAGGCTCTCTCAAGCAACTTACTGTCGAT 2501

RESULT 13
US-08-261-822A-1
Sequence 1, Application US/08261822A
Patent No. 5650553
GENERAL INFORMATION:
APPLICANT: Eckert, Joseph R. et al.
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5650553ris
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/261.822A
FILING DATE: 17-JUN-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6042 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-261-822A-1

Query Match 7.5%; Score 37.4; DB 1; Length 6042;
Best Local Similarity 58.6%; Pred. No. 0.33;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 74 TTCTTATATAACAGAGAGAAAACCTTGCCACACTGATGGGAGAGATTGAATGCG 133
Db 3609 TTGTTACATGAGAAATTAACAGCAAGTTTATTGAAAAGAGATGTTGAAGGGGTTTCATGGG 3668
Qy 134 AAACAGAGAAGGAGAGATGCACCTCAATAATATCTATTGCTGCCCAAT 184
Db 3669 AAACAGAGAAGGAGATGCACCAAGGCTCTCTCAAGCAACTTACTGTCGAT 3719

RESULT 14
PCT-US95-07744A-1
Sequence 1, Application PC/RUS9507744A
GENERAL INFORMATION:
APPLICANT: Trustees of The University of Pennsylvania
TITLE OF INVENTION: Plant Genes for Sensitivity to Ethylene
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/07744A
FILING DATE: 15-JUNE-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/261,822
FILING DATE: June 17, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6042 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US95-07744A-1

Query Match 7.5%; Score 37.4; DB 5; Length 6042;
Best Local Similarity 58.6%; Pred. No. 0.33;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 74 TTCTTATATAACAGAGAGAAAACCTTGCCACACTGATGGGAGAGATTGAATGCG 133
Db 3609 TTGTTACATGAGAAATTAACAGCAAGTTTATTGAAAAGAGATGTTGAAGGGGTTTCATGGG 3668
Qy 134 AAACAGAGAAGGAGAGATGCACCTCAATAATATCTATTGCTGCCCAAT 184
Db 3669 AAACAGAGAAGGAGATGCACCAAGGCTCTCTCAAGCAACTTACTGTCGAT 3719

RESULT 15
US-08-819-288-1
Sequence 1, Application US/08819288

Patent No. 5955652
GENERAL INFORMATION:
APPLICANT: Ecker, Joseph
APPLICANT: Alonso, Jose
TITLE OF INVENTION: PLANT GENES FOR SENSITIVITY TO ETHYLENE
TITLE OF INVENTION: AND PATHOGENS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5955652-rls
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/819,288
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Beardell, Lori Y.
REGISTRATION NUMBER: 34,293
REFERENCE/DOCKET NUMBER: UPN-2949
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-819-288-1

Query Match 7.5%; Score 37.4; DB 2; Length 6172;
Best Local Similarity 58.6%; Pred. No. 0.33;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 74 TTCTTATAACAGAGGAACTTGTGCCACACTGATGGCGAAGAGTTGATGCG 133
DB 3740 TTGTTACATGAGAGATTAACGCAAGTTTATTGAAAAGAGTGAAGGGGTTTCATGGG 3799
QY 134 AAACAGAGAGAGGAGCATGCACTCAATAATATCTATTGTCGCCAAT 184
DB 3800 AAACAGAGAGAGCTACCAAGAGCTGCTCTACAGCAACTTACTGTGCGAT 3850

Search completed: February 26, 2005, 20:48:31
Job time: 158 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: February 26, 2005, 19:05:25 ; Search time 499 Seconds
(without alignments)
5913.585 Million cell updates/sec

Title: US-10-621-901-26
Perfect score: 498
Sequence: 1 ggcctgcgaacagatgcctc.....aaattctgcgaagctgca 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
20: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	497	99.8	498	17	US-10-621-901-26
2	478.6	96.1	498	10	US-09-991-936-381
3	478.6	96.1	498	17	US-10-621-901-426
4	41.8	8.4	2172	15	US-10-154-419-9
5	41.8	8.4	2634	14	US-10-154-419-7
6	40.2	8.1	8866	15	US-10-239-676-140
7	40.2	8.1	8866	15	US-10-240-453-154
8	39.6	8.0	5862	15	US-10-311-455-518
9	38.8	7.8	1830121	17	US-10-329-670-1
10	38.8	7.8	1830121	18	US-10-158-865-1
11	38.4	7.7	15295	9	US-09-764-877-3404

12	38.4	7.7	15295	17	US-10-242-515-3404	Sequence 3404, Ap
13	37.6	7.6	3020	16	US-10-252-157-347	Sequence 347, App
14	37.4	7.5	443	9	US-09-770-444-856	Sequence 856, App
15	37.4	7.5	892	13	US-10-027-633-163122	Sequence 163122,
16	37.4	7.5	892	13	US-10-027-633-163122	Sequence 163123,
17	37.4	7.5	892	17	US-10-027-633-163122	Sequence 163122,
18	37.4	7.5	892	17	US-10-027-633-163122	Sequence 163123,
19	37.4	7.5	4746	16	US-10-385-521-10	Sequence 10, Appl
20	37.4	7.5	4746	16	US-10-385-521-10	Sequence 13, Appl
21	37.4	7.5	6022	16	US-10-385-521-11	Sequence 11, Appl
22	37.4	7.5	6022	18	US-10-602-475A-14	Sequence 14, Appl
23	37.2	7.5	2333	19	US-10-741-600-535	Sequence 535, App
24	37.2	7.5	2925	19	US-10-741-600-534	Sequence 534, App
25	37.2	7.5	2982	19	US-10-741-600-533	Sequence 533, App
26	37.2	7.5	3395	19	US-10-741-600-536	Sequence 536, App
27	36.2	7.3	952	17	US-10-424-599-20505	Sequence 20505, A
28	36.2	7.3	495269	17	US-10-398-221-8	Sequence 8, Appl1
29	36.2	7.3	301108	17	US-10-398-221-2058	Sequence 2058, Ap
30	35.6	7.1	1253	18	US-10-425-115-100766	Sequence 100766,
31	35.6	7.1	107745	18	US-10-322-281-268	Sequence 268, App
32	35.6	7.1	300000	15	US-10-262-552-33	Sequence 33, Appl
33	35.6	7.1	300000	18	US-10-703-210-33	Sequence 33, Appl
34	35.6	7.1	684707	17	US-10-398-221-9	Sequence 9, Appl1
35	35.4	7.1	269223	17	US-10-672-787-41	Sequence 41, Appl
36	34.8	7.0	798	13	US-10-027-633-124643	Sequence 124643,
37	34.8	7.0	798	17	US-10-027-633-124643	Sequence 124643,
38	34.8	7.0	200000	18	US-10-672-764A-11	Sequence 31, Appl
39	34.6	6.9	270	9	US-09-923-876-5308	Sequence 5308, Ap
40	34.6	6.9	270	9	US-09-923-876-5308	Sequence 5308, Ap
41	34.6	6.9	283	9	US-09-923-876-5178	Sequence 5178, Ap
42	34.6	6.9	283	10	US-09-923-876-5178	Sequence 5178, Ap
43	34.6	6.9	266	17	US-10-425-114-5475	Sequence 5475, Ap
44	34.6	6.9	3618	17	US-10-282-122A-35574	Sequence 35574, A
45	34.6	6.9	168575	16	US-10-178-194-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-621-901-26
; Sequence 26, Application US/10621901
; Publication No. US20040067516A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-8-C3
; CURRENT APPLICATION NUMBER: US/10/621, 901
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/319,414
; PRIOR FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 2313
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (403)..(403)
; OTHER INFORMATION: n = unknown at position 403
US-10-621-901-26

Query Match 99.8%; Score 497; DB 17; Length 498;
Best Local Similarity 100.0%; Pred. No. 3.6e-135;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGCCTGCAGACAGATGCCCTTCGACTCACCACCGAATATGAGCTGTCTACGATCAAAAT 60
|||||

Db 1 GGCCCTGGACAGATGCGCTTCCAGCTCAGACCGAATATGAGCTGTCTACATCAAAAT 60
Qy 61 GCCACCATGGAATGTTCTTAATTAACAGAGAGAGAAAATTGTGCCACACTGATGGCGAA 120
Db 61 GCCACCATGGAATGTTCTTAATTAACAGAGAGAGAAAATTGTGCCACACTGATGGCGAA 120
Qy 121 GGAGTTGAATGCGAACAAGAGAGAGAGATGCGACCTCAAAATATATCTATTTGCTGCC 180
Db 121 GGAGTTGAATGCGAACAAGAGAGAGAGATGCGACCTCAAAATATATCTATTTGCTGCC 180
Qy 181 CAATTAATTTCTGGAGTCGGTGGATCTTTATACAGCTTAGGTGATCCTATATGAGAC 240
Db 181 CAATTAATTTCTGGAGTCGGTGGATCTTTATACAGCTTAGGTGATCCTATATGAGAC 240
Qy 241 GATTAACACCAAAAATCTAATAACACAGACCTGATTAATTTCTTATTTCTTCGTATG 300
Db 241 GATTAACACCAAAAATCTAATAACACAGACCTGATTAATTTCTTATTTCTTCGTATG 300
Qy 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGCTGTCTTAAAGTTCAATTTGCCCA 360
Db 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGCTGTCTTAAAGTTCAATTTGCCCA 360
Qy 361 AGCTTGACGCCAACAAATGATTAATATGATCCTAGATGTTANGAGCATGGTGTAAAT 420
Db 361 AGCTTGACGCCAACAAATGATTAATATGATCCTAGATGTTANGAGCATGGTGTAAAT 420
Qy 421 TGGGTATACCTAGATCACTTTAATATTTTGTGCAAGTTGATTTGTTCTTAA 480
Db 421 TGGGTATACCTAGATCACTTTAATATTTTGTGCAAGTTGATTTGTTCTTAA 480
Qy 481 ATTCTGCCAAGAGCTGCA 498
Db 481 ATTCTGCCAAGAGCTGCA 498

RESULT 2

US-09-991-936-381
; Sequence 381, Application US/09991936
; Publication No. US20030073827A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Schluchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-6-C1
; CURRENT APPLICATION NUMBER: US/09/991,936
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US/09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1959
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 381
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-09-991-936-381

Query Match 96.1%; Score 478.6; DB 10; Length 498;
Best Local Similarity 98.8%; Pred. No. 8,9e-130;
Matches 492; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
Qy 1 GGCCCTGGACAGATGCGCTTCCAGCTCAGACCGAATATGAGCTGTCTACATCAAAAT 60
Db 1 GGCCCTGGACAGATGCGCTTCCAGCTCAGACCGAATATGAGCTGTCTACATCAAAAT 60
Qy 61 GCCACCATGGAATGTTCTTAATTAACAGAGAGAGAAAATTGTGCCACACTGATGGCGAA 120
Db 61 GCCACCATGGAATGTTCTTAATTAACAGAGAGAGAAAATTGTGCCACACTGATGGCGAA 120

Qy 121 GGAGTTGAATGCGAACAAGAGAGAGAGATGCGACCTCAAAATATATCTATTTGCTGCC 180
Db 121 GGAGTTGAATGCGAACAAGAGAGAGAGATGCGACCTCAAAATATATCTATTTGCTGCC 180
Qy 181 CAATTAATTTCTGGAGTCGGTGGATCTTTATACAGCTTAGGTGATCCTATATGAGAC 240
Db 181 CAATTAATTTCTGGAGTCGGTGGATCTTTATACAGCTTAGGTGATCCTATATGAGAC 240
Qy 241 GATTAACACCAAAAATCTAATAACACAGACCTGATTAATTTCTTATTTCTTCGTATG 300
Db 241 GATTAACACCAAAAATCTAATAACACAGACCTGATTAATTTCTTATTTCTTCGTATG 300
Qy 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGCTGTCTTAAAGTTCAATTTGCCCA 360
Db 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGCTGTCTTAAAGTTCAATTTGCCCA 360
Qy 361 AGCTTGACGCCAACAAATGATTAATATGATCCTAGATGTTANGAGCATGGTGTAAAT 420
Db 361 AGCTTGACGCCAACAAATGATTAATATGATCCTAGATGTTANGAGCATGGTGTAAAT 420
Qy 421 TGGGTATACCTAGATCACTTTAATATTTTGTGCAAGTTGATTTGTTCTTAA 480
Db 421 TGGGTATACCTAGATCACTTTAATATTTTGTGCAAGTTGATTTGTTCTTAA 480
Qy 481 ATTCTGCCAAGAGCTGCA 498
Db 481 A-TCTGCCAAGAGCTGCA 497

RESULT 3

US-10-621-901-426
; Sequence 426, Application US/10621901
; Publication No. US20040067516A1
; GENERAL INFORMATION:
; APPLICANT: Brandt, Kevin S.
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Schluchcomb, Dan T.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA HEAD, NERVE CORD, HINDGUT AND MALPIGHIAN TUBULE NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, PROTEINS AND USES THEREOF
; FILE REFERENCE: FC-8-C3
; CURRENT APPLICATION NUMBER: US/10/621,901
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/319,414
; PRIOR FILING DATE: 2002-07-22
; NUMBER OF SEQ ID NOS: 2313
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 426
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Ctenocephalides felis
US-10-621-901-426

Query Match 96.1%; Score 478.6; DB 17; Length 498;
Best Local Similarity 98.8%; Pred. No. 8,9e-130;
Matches 492; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
Qy 1 GGCCCTGGACAGATGCGCTTCCAGCTCAGACCGAATATGAGCTGTCTACATCAAAAT 60
Db 1 GGCCCTGGACAGATGCGCTTCCAGCTCAGACCGAATATGAGCTGTCTACATCAAAAT 60
Qy 61 GCCACCATGGAATGTTCTTAATTAACAGAGAGAGAAAATTGTGCCACACTGATGGCGAA 120
Db 61 GCCACCATGGAATGTTCTTAATTAACAGAGAGAGAAAATTGTGCCACACTGATGGCGAA 120
Qy 121 GGAGTTGAATGCGAACAAGAGAGAGAGATGCGACCTCAAAATATATCTATTTGCTGCC 180
Db 121 GGAGTTGAATGCGAACAAGAGAGAGAGATGCGACCTCAAAATATATCTATTTGCTGCC 180
Qy 181 CAATTAATTTCTGGAGTCGGTGGATCTTTATACAGCTTAGGTGATCCTATATGAGAC 240
Db 181 CAATTAATTTCTGGAGTCGGTGGATCTTTATACAGCTTAGGTGATCCTATATGAGAC 240

QY 241 GATPAACCAAAAAATCTAAAAACACGACCTGATAGTTTCTTATTTCTTCTGATG 300
| | | | |
Db 241 GATPAACCAAAAAATCTAAAAACACGACCTGATAGTTTCTTATTTCTTCTGATG 300
QY 301 CTGGAGCTTCAACCGGTTATGCTTGGCTAGCGTCTGTCTTAAAGTTCTACATTTGGCCA 360
| | | | |
Db 301 CTGGAGCTTCAACCGGTTATGCTTGGCTAGCGTCTGTCTTAAAGTTCTACATTTGGCCA 360
QY 361 ACCTTGACGCCAACATTTGATATGATGCTGATGTTTANGAGCATGGTGGTAAAGT 420
| | | | |
Db 361 ACCTTGACGCCAACATTTGATATGATGCTGATGTTTANGAGCATGGTGGTAAAGT 420
QY 421 TGGGTGATCTAGAGTCACTTTATATTTTTCGCAAGTTGATGATTTGTTCTTAA 480
| | | | |
Db 421 TGGGTGATCTAGAGTCACTTTATATTTTTCGCAAGTTGATGATTTGTTCTTAA 480
QY 481 ATTCTGCCAAGAGCTGCA 498
| | | | |
Db 481 A-TCTGCCAAGAGCTGCA 497

RESULT 4
US-10-154-419-9
; Sequence 9, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory A.J.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 9
; LENGTH: 2172
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2172)
US-10-154-419-9

Query Match 8.4%; Score 41.8; DB 15; Length 2172;
Best Local Similarity 47.8%; Pred. No. 0.26; Mismatches 163; Indels 6; Gaps 1;
Matches 155; Conservative 0; Mismatches 163; Indels 6; Gaps 1;
QY 181 CAATTAATTTCTGAGTGGTGATCTTTATATCTACATTTAGTGATCTTATATGAGAC 240
| | | | |
Db 706 CAATTAATTTCTGAGTGGTGAGGAGGAGAACTCTTTATATCTGGAACAGCTTTCTGAT 765
QY 241 GATPAACCAAAAAATCTAAAAACACGACCTGATAGTTTCTTATTTCTTCTGATG 300
| | | | |
Db 766 GATTTCTGTGCCACACCAAGTCTTCTCTATATAGAACCGGTTTATGCTATGCAATC 825
QY 301 CTGGAGCTTCAACCGGTTATGCTTGGCTAGCGTCTGTCTTAAAGTTCTACATTTGGCCA 360
| | | | |
Db 826 TTAGGCGCTGCTATTTGGCTATGTTTGGAGAGAACCTGTAACCATATACATTTGATGTT 885
QY 361 ACCTTGACGCCAACATTTGATATGATGCTGATGTTTANGAGCATGGTGGTGG 414
| | | | |
Db 886 GCTATGGAGAAAGACCTGATGTCATGAGATGATCGCGATGTTGGAGCTTGGTGG 945
QY 415 TTAAGTTGGGTGATCTAGATCACTTTATATTTTTCGCAAGTTGATGATTTGTT 474
| | | | |
Db 946 ATTGGGTTTCTTCTATCATGATCTTTGCTTGTCTTAAATATACCTTTTCTTGGCTTT 1005
QY 475 CCTAAATTTCTGCCAAGAGCTGCA 498
| | | | |

Db 1006 CCAAAACATTTACAGGTACGACGA 1029
| | | | |

RESULT 5
US-10-154-419-7
; Sequence 7, Application US/10154419
; Publication No. US20030143675A1
; GENERAL INFORMATION:
; APPLICANT: Curtiss, Rory A.J.
; APPLICANT: Glucksmann, Maria Alexandra
; APPLICANT: Meyers, Rachel E.
; TITLE OF INVENTION: NOVEL 38594, 57312, 53659, 57250, 63760, 49938, 32146,
; TITLE OF INVENTION: 57259, 67118, 67067, 62092, 8099, 46455, 54414, 53763,
; TITLE OF INVENTION: 67076, 67102, 44181, 67084FL, 67084ALT, FBH58295FL, 57255,
; TITLE OF INVENTION: AND 57255alt MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-249
; CURRENT APPLICATION NUMBER: US/10/154,419
; CURRENT FILING DATE: 2002-05-22
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 7
; LENGTH: 2634
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (104)...(2275)
US-10-154-419-7

Query Match 8.4%; Score 41.8; DB 15; Length 2634;
Best Local Similarity 47.8%; Pred. No. 0.29; Mismatches 163; Indels 6; Gaps 1;
Matches 155; Conservative 0; Mismatches 163; Indels 6; Gaps 1;
QY 181 CAATTAATTTCTGAGTGGTGATCTTTATATCTACATTTAGTGATCTTATATGAGAC 240
| | | | |
Db 809 CAATTAATTTCTGAGTGGTGAGGAGGAGAACTCTTTATATCTGGAACAGCTTTCTGAT 868
QY 241 GATPAACCAAAAAATCTAAAAACACGACCTGATAGTTTCTTATTTCTTCTGATG 300
| | | | |
Db 869 GATTTCTGTGCCACACCAAGTCTTCTCTATATAGAACCGGTTTATGCTATGCAATC 928
QY 301 CTGGAGCTTCAACCGGTTATGCTTGGCTAGCGTCTGTCTTAAAGTTCTACATTTGGCCA 360
| | | | |
Db 929 TTAGGCGCTGCTATTTGGCTATGTTTGGAGAGAACCTGTAACCATATACATTTGATGTT 988
QY 361 ACCTTGACGCCAACATTTGATATGATGCTGATGTTTANGAGCATGGTGGTGG 414
| | | | |
Db 989 GCTATGGAGAAAGACCTGATGTCATGAGATGATCGCGATGTTGGAGCTTGGTGG 1048
QY 415 TTAAGTTGGGTGATCTAGATCACTTTATATTTTTCGCAAGTTGATGATTTGTT 474
| | | | |
Db 1049 ATTGGGTTTCTTCTATCATGATCTTTGCTTGTCTTAAATATACCTTTTCTTGGCTTT 1108
QY 475 CCTAAATTTCTGCCAAGAGCTGCA 498
| | | | |
Db 1109 CCAAAACATTTACAGGTACGACGA 1132
| | | | |
RESULT 6
US-10-239-676-140
; Sequence 140, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968

Qy 436 TCACTTAATATTTTGCACGTTGATTTGCTTAATTCG 486
Db 2932 TTGGAATTTGATTTTGGTAATGTTAGTTTATTAATGTGTG 2982

RESULT 9
US-10-329-670-1/c
; Sequence 1, Application US/10329670
; Publication No. US2004001850A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE REFERENCE: P186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature

```
LOCATION: (120038) .. (120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344) .. (121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167) .. (122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336) .. (122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360) .. (131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910) .. (139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398) .. (140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750) .. (142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058) .. (145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171) .. (145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942) .. (145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197) .. (147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841) .. (150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500) .. (152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530) .. (152530)
OTHER INFORMATION: n equals a, t, g or c

Query Match 7.8%; Score 38.8; DB 17; Length 1830121;
Best Local Similarity 49.3%; Pred. No. 58;
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy 262 ACACCGACATGATAGTTTCTTATTTCTTCTGATGCTCGGACCTGCAACCGGTTAT 321
Db 1506177 AAGTGTAGATTATTAAGCCTTGATTTATGCTGAGAAATTAACGCAAGAGTTAT 1506118

Qy 322 GCCTGGCTAGCGCTGTCTTAAGTTCTACATTTGCGCAAGTGAAGCCAAATGAT 381
Db 1506117 AGCTTGATGCATTAATATTAACGAATGTGCTGAGCATTAAGCGAAGAAAGCATTTGCG 1506058
```

```
Qy 382 AATAATGATCTTGATGCTTANGACATGTGTTAAGTTGGGTGATAGATCAACT 441
Db 1506057 AGTTTATGCTGTAAGCGTTGAGGGAATGGAAGGCGATTAAGCTGTGCTTAAGCAACT 1505998

Qy 442 TTAATATTTTTCGAACGTTGAT 464
Db 1505997 TTAATACTTTTGAAGATTAAGAT 1505975

RESULT 10
US-10-158-865-1/c
Sequence 1, Application US/10158865
Publication No. US20040203093A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Fragment
FILE REFERENCE: PB16P2C1D1
CURRENT APPLICATION NUMBER: US/10/158, 865
CURRENT FILING DATE: 2002-06-03
PRIORITY APPLICATION NUMBER: US 09/557,884
PRIORITY FILING DATE: 2000-04-25
PRIORITY APPLICATION NUMBER: US 08/476,102
PRIORITY FILING DATE: 1995-06-07
PRIORITY APPLICATION NUMBER: US 08/426,787
PRIORITY FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4747) .. (4747)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921) .. (9921)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10150) .. (10150)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298) .. (29298)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543) .. (36543)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551) .. (36551)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636) .. (36636)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40810) .. (40810)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416) .. (44416)
OTHER INFORMATION: n equals a, t, c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905) .. (44905)
OTHER INFORMATION: n equals a, t, c, or g
```

NAME/KEY: misc_feature	LOCATION: (119924) ..(119924)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (120038) ..(120038)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (121344) ..(121344)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (122167) ..(122167)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (122336) ..(122336)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (131340) ..(131340)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (139510) ..(139510)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (140398) ..(140398)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (142750) ..(142750)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (145058) ..(145058)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (145711) ..(145711)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (147197) ..(147197)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (150841) ..(150841)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (152500) ..(152500)	OTHER INFORMATION: n equals a,t,c, or g
NAME/KEY: misc_feature	LOCATION: (152530) ..(152530)	OTHER INFORMATION: n equals a,t,c, or g

	Query Match	7.8%	Score 36.9	DB 18;	Length 1830121.
	Best Local Similarity	49.3%	Pred. No. 58;		
	Matches 100; Conservative	0;	Mismatches 103;	Indels 0;	Gaps 0.
Oy	262 ACACCACACGATAGTTTTTCTATTTTTCGTGTAAGTCGGACCGCAACGGTAT				

Db	1506177	AAAGTAGAGTTATTTAAAGCCCTTTGATTTTATGGGTGAGAAATTACAGCAAGAAAGTTAT	1506118
Qy	322	GCCTTGGCTAGCGCTGTCGTTAAAGTTCTACATTTGGCCACGTTGAGCCCAACTGAT	381
Db	1506117	AGCTTGATGCACTTAATATATTAACGAATTTTGCGATTTAGAGCATTAAGGAAAGCAATTTGGCG	1506058
Qy	382	AATATATGATCTCTAATGTTTANGAGCATGCGTGTAACTTGGTGGATATCTATGATCAACT	441
Db	1506057	AGTTTATGTGTGTTAAAGGGTTTANGAGAAATGAGAAAGGCGATTAAGCCTGTGCTTAAGAACT	1505988
Qy	442	TTAATATTTTTCGACGTTGAT	464
Db	1505997	TTAAAGCTTTTGAAGATAAGAT	1505975

```

RESULT 11
US-09-764-877-3404
: Sequence 3404, Application US/09764877
: Patent No. US20020147740A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.,
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC005
: CURRENT APPLICATION NUMBER: US/09/764,877
: CURRENT FILING DATE: 2001-01-17
: Prior application data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 4031
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3404
: LENGTH: 15295
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-764-877-3404

```

Query Match	7.7%;	Score 38.4;	DB 9;	Length 15295;
Best Local Similarity	60.0%;	Pred. No. 7;		
Matches 63;	Conservative 0;	Mismatches 42;	Indels 0;	

Qy	Db
374	12389
CAATTGATTAATATGATCCTAATGGTTANGAGCATGGTGGTTAAFTGGGTGATCTAG	CAATTGATTGAAAGATTACTGTGAGGATTTGACTTATGCTTTTGAATAGAAAATTATATA
433	12448

Oy 434 GATCACTTTAATTTTTCGAACGTTGATGGATTGTTCCTA 478
 |||||
 Db 1249 GAACAAGTAACTACTTTTAAAGCCTTATTAGATGTTATCTA 1244

RESULT 12
US-10-242-515-3404

Sequence 3404, Application US/10242515
Publication No. US20040009488A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.,
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005051
CURRENT APPLICATION NUMBER: US/10/242,515
CURRENT FILING DATE: 2002-09-13
PRIOR APPLICATION NUMBER: 09/764,877
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,436
PRIOR FILING DATE: 2000-07-11

```

? PRIOR APPLICATION NUMBER: 60/225,447
? PRIOR FILING DATE: 2000-08-14
? PRIOR APPLICATION NUMBER: 60/218,290
? PRIOR FILING DATE: 2000-07-14
? Remaining prior Application data removed - See File Wrapper or PALM
? NUMBER OF SEQ ID NOS: 4031
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 3404
? LENGTH: 15295
? TYPE: DNA
? ORGANISM: Homo sapiens
?
US-10-242-515-3404

```

Query Match	7.7%	Score 38.4	DB 17	Length 15295
Best Local Similarity	60.0%	Pred. No. 7		
Matches 63	Conservative 0	Mismatches 42	Indels 0	Gaps 0

QY 374 CAATTGATTAATATGATCCCTAGATGGTTANGAGCATGGTGGTTAAGTTGGGTACTACG 433

Db 12389 CAATTGTTAGAAAGATACTGCTAGAGATTGACTTATGCTTTTTCGATAGAAAATTATATA 12448

Qy	434	GATCACTTAAATATTTTGGCAAGCTGATGGATGTTCCCTA	478
Db	12449	GACACGTTAACTATTTTAAAGCTTATTAGATGTTATCTA	12493

RESULT 13
US-10-252-157-347

; Sequence 347, Application US/10252157
; Publication No. US20030190640A1

; GENERAL INFORMATION:

APPLICANT: Faris, Mary
APPLICANT: Pearson, Cecelia I.

TITLE OF INVENTION: GENES EXP

FILE REFERENCE: PA-0027-1 US
CURRENT APPLICATION NUMBER: US/10/252,157

CURRENT FILING DATE: 2002-10-01

PRIOR APPLICATION NUMBER: 60/295,048

; PRIOR FILING DATE: 2001-05-31
 ; NUMBER OF SEQ ID NOS: 501

```

; SOFTWARE: PERL Program
; SEO ID NO 347

```

LENGTH: 3020

```

/
;
; TYPE: DNA
; ORGANISM: Homo sapiens

```

FEATURE: HOW BAYREITH

```

NAME/KEY: misc feature
OTHER INFORMATION: Incvte ID NO IIS20030190640A1 372313.6

```

OTHER INFORMATION: Incyte ID No. US20030190640A1 372313.6
US-10-252-157-347

Query Match	7.6%	Score 37.6	DB 16	Length 3020
Best Local Similarity	47.6%	Pred. No. 5.3		
Matches 152, Conservative	0	Mismatches	155	Indels 12, Gaps 1

QY 192 TGAAGTCGGTGAATCTTTATACACACTTAGGCGTATCCTAATGAGACGATACACCA 251
 557 TGAATGGGTGAACCTCCCATCTGCGCTTTGGGTAATTTCTTAATAGAGATTTTGCCA 616
 Db

QY 252 AAATCTAAACACGAGCACTGATTAAGTTTTCTTATTTCTTCGTAGCTCGAACCTGC 311
 617 ATTGGAATTCCTTATATATATTGGCGCTGTAGAAACAGGACTATTAATGGCTCTTT 676

QY 312 AACCGTTATGCTGGCTAGCGCTGCTCTAAAGTTTCATACATTGGCCAAAGTTGACGCC 371
 Db 677 GATTGACCTTTGTTGGCATCTCTGCGCAATGTTTATGTTCACATCGATTGTGTAA 736

QY 312 AACATTTGATAATA-----TGATCTGAGATGGTTANAGACATGGGTGTTAG 419

Db 737 CACGATGATCTGATCATTAACCTCCCACTGACATCTGTTGGGTCGGTCATGGTCTTGG 796

QY 420 TTGGTATTACTAGCATCACTTAAATTTTGGCAAGTTAGATTTGTTCCCA 479
Db 797 CTTCTGATTTGTGACAGAGTAAAGTGTCACTGCCATTCCTTTTTCTTTTGGCCCA 856

QY 480 AATTCGCAAGAGCTGCA 498
DB 857 CACACTTCCAAAGAGCA 875

RESULT 14

US-09-770-444-856
Sequence 856, Application US/09770444
Patent No. US20020023280A1
GENERAL INFORMATION:
APPLICANT: Gorlach, Jorn
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy W.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Matthew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Moesener, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Kricker, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2027 (PARA-016PRV)
CURRENT APPLICATION NUMBER: US/09/770,444
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/178,502
PRIOR FILING DATE: 2000-01-27
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 856
LENGTH: 443
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-770-444-856

Query Match 7.5%; Score 37.4; DB 9; Length 443;
Best Local Similarity 58.6%; Pred. No. 2.3;
Matches 65; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 74 TTCTTATAAAGAGAGGAACTTGGCCACACTGATGGGAGAGGATGATGCG 133
DB 270 TTGTTACCAAGAGAAATACGCAAGTTTATTGAAAGAGATGTTGAGGGGTTTCATGGG 329
QY 134 AAACAGAGAGAGGAGCATGCACTCAATAATATATCTATTGCTGCCCAAT 184
DB 330 AAACAGAGAGAGGAGGCTACCAAGGCTGCTCTCAAGCACTTACTGTGCGAT 380

RESULT 15

US-10-027-632-163122/C
Sequence 163122, Application US/10027632
Publication No. US2002019837A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 163122
LENGTH: 892
TYPE: DNA
ORGANISM: Human
US-10-027-632-163122

Query Match 7.5%; Score 37.4; DB 13; Length 892;
Best Local Similarity 50.6%; Pred. No. 3.3;
Matches 89; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 319 TATGCTTGCTAGCGTCTGTAAAGTTCTACATTTGCCCAAGTGAAGCCCAAT 378
DB 736 TAGGTATGCTTCTTATTCTAATTTGTAATTCAGAGGAAATCAACAAATGTC 677
QY 379 GATATATGATCCTGATGATGTTANGAGCATGGTGTAGTGGTACTAGATCA 438
DB 676 TAAATATATACGATATGCGATATGAGCACCAAAATTAATGTTGAAATTTCTATATGTC 617
QY 439 ACTTATATTTTGGCAAGTGTGATTTGTTCTTAAATTTGCGCAAGGC 494
DB 616 AGTAAATCATATGCAAGTGTGATTTCTGATGATATATTCACAAATCTGC 561

Search completed: February 26, 2005, 20:57:04
Job time: 510 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 18:45:12 ; Search time 3109 Seconds
(without alignments)
6097.139 Million cell updates/sec

Title: US-10-621-901-26
Perfect score: 498
Sequence: 1 ggcctcgagcaagatgcgcct.....aaattctgcaagagctgca 498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479086

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: gb_ests1:*
2: gb_ests2:*
3: gb_ests3:*
4: gb_ests4:*
5: gb_ests5:*
6: gb_ests6:*
7: gb_ests7:*
8: gb_ests8:*
9: gb_ests9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	497	99.8	498	2	BF731833 3085-57 h
2	478.6	96.1	498	4	BM056367 2092-22 h
3	147.8	29.7	993	3	BM056367 2092-22 h
4	109.2	21.9	920	8	CC131868 NDL.93017
5	107.6	21.6	398	5	BK553516 BK553516
6	73.8	14.8	444	7	CO340715 EP11205.3
7	72	14.5	412	4	BT513920 BT160014A
8	71.2	14.3	371	4	BT513842 BT160014A
9	69.2	13.9	632	5	BM618196 170006592
10	66	13.3	561	5	BK551134 BK551134
11	50.8	10.2	463	1	AU209481 AU209481
12	46.6	9.4	605	5	BU214734 603755346
13	43	8.6	484	1	AU202605 AU202605
14	42.8	8.6	840	9	CL865155 CL865155
15	42.2	8.5	555	7	CO715174 DG14-245n
16	41	8.2	707	1	AJ741950 AJ741950
17	40.8	8.2	1021	9	CNS01681 CNS01681
18	40.6	8.2	802	8	BZ400007 BZ400007
19	40.6	8.2	838	5	BZ393310 E1NAG61TR
20	40.4	8.1	564	5	BK517949 BK517949
21	40.4	8.1	693	1	A1789248 A1789248
22	40.2	8.1	599	9	CPA56299 CPA56299
23	39.8	8.0	1101	9	CNS00782 CNS00782
24	39.6	8.0	329	7	CK986460 4088369 B

25	39.6	8.0	521	1	AA043744
26	39.6	8.0	725	1	AJ741939
27	39.6	8.0	760	5	BU199773 604157396
28	39.6	8.0	1101	9	CNS006PN
29	39.4	7.9	632	9	CNS0101N3
30	39.4	7.9	765	8	BZ777990
31	39.2	7.9	592	1	A1663317
32	39.2	7.9	673	1	A1788287
33	38.8	7.8	685	8	BZ023383
34	38.6	7.8	539	8	A0003681
35	38.6	7.8	654	8	CC167112
36	38.6	7.8	2394	3	AK049253
37	38.6	7.8	3412	3	AK049944
38	38.4	7.7	645	8	AZ524922
39	38.4	7.7	851	2	AW045107
40	38.4	7.7	1101	9	CNS0100X
41	38.2	7.7	607	7	CN433293
42	38.2	7.7	916	9	CC971176
43	38	7.6	600	9	CNS07586
44	37.8	7.6	456	8	BZ616865
45	37.8	7.6	642	8	CC163691

ALIGNMENTS

RESULT 1
BF731833
LOCUS
DEFINITION
3085-57 hindgut and Malpighian tubule subcloned cDNA library
ACCESSION
BF731833
VERSION
BF731833.1 GI:22038812
KEYWORDS
EST.
SOURCE
ORGANISM
Ctenocephalides felis (cat flea)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Siphonaptera; Pulicidae; Pulicidae;
Ctenocephalides.
1 (bases 1 to 498)
Gaines, P.J., Brandt, K.S., Eisele, A.M., Wagner, W.P., Bozic, C.M. and
Wisniewski, N.
Analysis of expressed sequence tags from subcloned and
unsubcloned Ctenocephalides felis hindgut and Malpighian tubule
cDNA libraries.
J. Biol. Chem. 274(4):299-306 (2002)

JOURNAL
MEDLINE
PUBMED
COMMENT
Contact: Gaines, Patrick J.
Pharmaceutical Discovery Group
Heska Corporation
1613 Prospect Parkway, Fort Collins, CO 80525 USA
Tel: 970 493-7272
Fax: 970 472-1644
Email: gaines@heska.com
Seq primer: CGA GCG ATC CAC TAG.
Location/Qualifiers
1. 498
/organism="Ctenocephalides felis"
/mol_type="mRNA"
/db_xref="taxon:7515"
/clone="3085-57"
/sex="Female and male (4:1 ratio)"
/tissue_type="hindgut and Malpighian tubule"
/dev_stage="unfed adult and 24 hour cat blood-fed adults (1:1 ratio)"
/clone_lib="hindgut and Malpighian tubule subcloned cDNA library"
/note="The library was made using the PCR-select (CDN
Subtraction Kit (CDNTECH Laboratories, Inc.) with 2 (g
hindgut and Malpighian tubule mRNA used as the
'tester' cDNA and 2 (g carcass (all flea tissues minus
the hindgut and Malpighian tubules) to make the 'driver'

FEATURES

source

cDNA. Suppression PCR and subtractive hybridization techniques are then used to enrich for cDNAs in the tester pool that are not also present in the driver pool.

ORIGIN

Query Match 99.8%; Score 497; DB 2; Length 498;

Best Local Similarity 100.0%; Pred. No. 1.3e-131;

Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GGCCCTGGACCAAGATGCGCTTGACCTCACCGAATATGAGCTGCTACGATCAAAAT 60
DB 1 GGCCCTGGACCAAGATGCGCTTGACCTCACCGAATATGAGCTGCTACGATCAAAAT 60
QY 61 GCCCATTGAAGTCTTATATAACAGAAAGAGAAAACTTTGGCCACACTGATGGCGAA 120
DB 61 GCCCATTGAAGTCTTATATAACAGAAAGAGAAAACTTTGGCCACACTGATGGCGAA 120
QY 121 GGAGTGAATGCGAAACAGAAAGAGAGCATGCGACCTCAATATATCTATTGCTGCC 180
DB 121 GGAGTGAATGCGAAACAGAAAGAGAGCATGCGACCTCAATATATCTATTGCTGCC 180
QY 181 CAATTAATTTCTGAGTCGGTGGATCTTTATCTACACTTATAGTATCTATATGAGAC 240
DB 181 CAATTAATTTCTGAGTCGGTGGATCTTTATCTACACTTATAGTATCTATATGAGAC 240
QY 241 GATTAACACCAAAAAATCTAAACACGACGACATGATTAATTTCTTATTTCTTGTATG 300
DB 241 GATTAACACCAAAAAATCTAAACACGACGACATGATTAATTTCTTATTTCTTGTATG 300
QY 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGGTCTGTCTAAAGTTCTACATTTGCCA 360
DB 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGGTCTGTCTAAAGTTCTACATTTGCCA 360
QY 361 ACCTTACCGCAACATGTAATTAATGATCTAGATGTTTANGAGATGCTGTTAAGT 420
DB 361 ACCTTACCGCAACATGTAATTAATGATCTAGATGTTTANGAGATGCTGTTAAGT 420
QY 421 TGGGTACTACTAGATCACTTTAATATTTTGGCAACGTTGATGTTTCTCTAAA 480
DB 421 TGGGTACTACTAGATCACTTTAATATTTTGGCAACGTTGATGTTTCTCTAAA 480
QY 481 ATTTCGCCAAGAGCTGCA 498
DB 481 ATTTCGCCAAGAGCTGCA 498

RESULT 2
BM056367 498 bp mRNA linear EST 04-SEP-2002
LOCUS 2092-22 hindgut and Malpighian tubule subtracted cDNA library
DEFINITION Ctenocephalides felis cDNA clone 2092-22, mRNA sequence.
ACCESSION BM056367
VERSION BM056367.1 GI:16898176
KEYWORDS EST.
SOURCE Ctenocephalides felis (cat flea)
ORGANISM Ctenocephalides felis
Fukayaoka; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Siphonaptera; Pulicidae; Pulicines;
Ctenocephalides.
1 (bases 1 to 498)
Gaines,P.J., Brandt,K.S., Eisele,A.M., Wagner,W.P., Bozic,C.M. and
Wisniewski,N.
REFERENCE
AUTHORS
TITLE
ANALYSIS of expressed sequence tags from subtracted and
unsubtracted Ctenocephalides felis hindgut and Malpighian tubule
cDNA libraries
JOURNAL Insect Mol. Biol. 11 (4), 299-306 (2002)
MEDLINE 22140248
PUBMED 12144694
COMMENT Contact: Gaines, Patrick J.
Pharmaceutical Discovery Group
Heska Corporation
1613 Prospect Parkway, Fort Collins, CO 80525 USA
```

Tel: 970 493-7272
Fax: 970 472-1644
Email: gaines@heska.com
Seq primer: CGA GCT CGC ATC CAC TAC.
Location/Qualifiers

FEATURES

source

1..498

/organism="Ctenocephalides felis"

/mol_type="mRNA"

/db_xref="taxon:7515"

/clone="2092-22"

/sex="female and male (4:1 ratio)"

/tissue_type="hindgut and Malpighian tubule"

/dev_stage="unfed adult and 24 hour cat blood-fed adults (1:1 ratio)"

/clone_lib="hindgut and Malpighian tubule subtracted cDNA library"

/note="The library was made using the PCR-select (cDNA subtraction kit (CLONTECH Laboratories, Inc.) with 2 (g hindgut and Malpighian tubule mRNA used as to make the 'tester' cDNA and 2 (g carcasses (all flea tissues minus the hindgut and Malpighian tubules) to make the 'driver' cDNA. Suppression PCR and subtractive hybridization techniques are then used to enrich for cDNAs in the tester pool that are not also present in the driver pool."

ORIGIN

Query Match 96.1%; Score 478.6; DB 4; Length 498;

Best Local Similarity 98.8%; Pred. No. 2.5e-126;

Matches 492; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
QY 1 GGCCCTGGACCAAGATGCGCTTGACCTCACCGAATATGAGCTGCTACGATCAAAAT 60
DB 1 GGCCCTGGACCAAGATGCGCTTGACCTCACCGAATATGAGCTGCTACGATCAAAAT 60
QY 61 GCCCATTGAAGTCTTATATAACAGAAAGAGAAAACTTTGGCCACACTGATGGCGAA 120
DB 61 GCCCATTGAAGTCTTATATAACAGAAAGAGAAAACTTTGGCCACACTGATGGCGAA 120
QY 121 GGAGTGAATGCGAAACAGAAAGAGAGCATGCGACCTCAATATATCTATTGCTGCC 180
DB 121 GGAGTGAATGCGAAACAGAAAGAGAGCATGCGACCTCAATATATCTATTGCTGCC 180
QY 181 CAATTAATTTCTGAGTCGGTGGATCTTTATCTACACTTATAGTATCTATATGAGAC 240
DB 181 CAATTAATTTCTGAGTCGGTGGATCTTTATCTACACTTATAGTATCTATATGAGAC 240
QY 241 GATTAACACCAAAAAATCTAAACACGACGACATGATTAATTTCTTATTTCTTGTATG 300
DB 241 GATTAACACCAAAAAATCTAAACACGACGACATGATTAATTTCTTATTTCTTGTATG 300
QY 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGGTCTGTCTAAAGTTCTACATTTGCCA 360
DB 301 CTGGAACCTGCAACCGGTTATGCTTGGCTAGCGGTCTGTCTAAAGTTCTACATTTGCCA 360
QY 361 ACCTTACCGCAACATGTAATTAATGATCTAGATGTTTANGAGATGCTGTTAAGT 420
DB 361 ACCTTACCGCAACATGTAATTAATGATCTAGATGTTTANGAGATGCTGTTAAGT 420
QY 421 TGGGTACTACTAGATCACTTTAATATTTTGGCAACGTTGATGTTTCTCTAAA 480
DB 421 TGGGTACTACTAGATCACTTTAATATTTTGGCAACGTTGATGTTTCTCTAAA 480
QY 481 ATTTCGCCAAGAGCTGCA 498
DB 481 ATTTCGCCAAGAGCTGCA 497
```

RESULT 3
CNS08HM7 993 bp mRNA linear HTC 07-JAN-2003
LOCUS Single read from an extremity of a full-length cDNA clone made from
DEFINITION Anopheles gambiae total adult females. 5-PRIME end of clone

FK0AA19CA02 of strain 6-9 of *Anopheles gambiae* (African malaria mosquito).
 BX012251
 VERSION BX012251.1 GI:27561471
 KEYWORDS HTC.
 SOURCE *Anopheles gambiae* (African malaria mosquito)
 ORGANISM *Anopheles gambiae*
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;

REFERENCE
 1 (bases 1 to 993)
 Anopheles.

AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (06-JAN-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

FEATURES
 source
 1..993
 Location/Qualifiers
 /organism="Anopheles gambiae"
 /mol_type="mRNA"
 /strain="6-9"
 /db_xref="taxon:7165"
 /clone="FK0AA19CA02"
 /plasmid="pME18S-FL"
 /note="end : 5-PRIME"

ORIGIN

Query Match 29.7%; Score 147.8; DB 3; Length 993;

Best Local Similarity 62.7%; Pred. No. 3.2e-31;
 Matches 262; Conservative 0; Mismatches 153; Indels 3; Gaps 2;

1 GGGCCGGAACAATGCGCTTCGACCTACACCGAATATGAGTGTCTACGATCAAAAT 60
 577 GGGCCGGAAGAGCGCCCTCTCTCCCTGACCACTGATGACGTGC-ACGTACGATGCGAAC 635
 61 GCCACCATTTGATGATCTTAAATTAACAGAGAGAGAACTTTGTGCCACACTGATGGCAA 120
 636 CACACGCGGAGATGATGAGAGCCGAGAGAGCCAAATCTCTGCGCAGAAATGCCACT 695
 121 GGAATT-CAATCGAAGACAGAGAGAGAGATCGACCTCAATTAATCTATTTCTGT 178
 696 CCGGCGCGGAGAAATGTGAACGAGAGAGAGACCTGCGCACACAGATCGTGTGTTATCG 755
 179 CCCAATTAATTTCTGAGTGGTGGATCTTATTAATCACTTAAGTATGATCCATATGG 238
 756 CCCAGTTATCTCCGCGATTTGAAGCTGCTCTACTACACGCTGGGCGTCTCGTACATGG 815
 239 ACGATAACACCAAAAATCTAAACACACGACATGATAGTTTCTTATTTCTTCTGTA 298
 816 ACGATAACATCAAGAGATCGAAGAACCGGCGCTGTCAGTGTCTGTAATTTCTGCGCA 875
 299 TGGTCGACCTGCAACCGGTTATGCTTGGCTAGCGTGTCTTAAAGTTCTACATTTGCC 358
 876 TGGTCGACCGGCGATGAGGTACACGCTGGCGTCTCTTCTGAACTGACATCTCGC 935
 359 CAAGTTGAGCGCAACATGATATATGATCTTATGATTTTANGAGCATGGTGT 416
 936 CTTGATGACGCGAGCATGACGATACGACATCCGCGGCTGCTGGTCTGGTGGAT 993

RESULT 4
 LOCUS CC131868 920 bp DNA linear GSS 16-APR-2003
 DEFINITION ND1_93017.T7 Notre Dame Liverpool Aedes aegypti genomic clone
 ACCESSION CC131868
 VERSION CC131868.1 GI:30000923
 KEYWORDS GSS
 SOURCE Aedes aegypti (yellow fever mosquito)
 ORGANISM Aedes aegypti
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Aedes; Stegomyia.

REFERENCE
 1 (bases 1 to 920)
 AUTHORS Loftus,B., Shetty,J., Knudson,D. and Severson,D.
 TITLE BAC end sequencing of *Aedes aegypti*
 JOURNAL Unpublished (2003)
 COMMENT Other_GSSs: ND1.93017.SP6
 Contact: Brendan Loftus
 Department of Eukaryotic Genomics
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-3543
 Fax: 301-838-0208
 Email: entae@tigr.org
 Library was provided by David Severson
 Seq primer: T7
 Class: BAC ends.

FEATURES
 source
 1..920
 Location/Qualifiers
 /organism="Aedes aegypti"
 /mol_type="genomic DNA"
 /strain="Liverpool"
 /db_xref="taxon:7159"
 /clone="ND1.93017"
 /clone_1lb="Notre Dame Liverpool"
 /note="Vector: pBCBAC1; Site 1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"

ORIGIN

Query Match 21.9%; Score 109.2; DB 8; Length 920;

Best Local Similarity 64.2%; Pred. No. 3.9e-20;
 Matches 181; Conservative 0; Mismatches 98; Indels 3; Gaps 1;

1 GGGCCGGAACAATGCGCTTCGACCTACACCGAATATGAGTGTCTACGATCAAAAT 60
 521 GATCCGGAAGAGCGCCCTTCTGTTAACTACGAGATGGGCGACATGACGAGAT 580
 61 GCCACCATTTGATGATCTTAAATTAACAGAGAGAGAACTTTGTGCCACACTGATGGCAA 119
 581 CAAGCTAAGAGATGTTTGAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 640
 120 -AGAGATTGAATGCGAAG 177
 641 CAGGAGACTAATGTGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 700
 178 GCCAATTAATTTCTGAGTGGTGGATCTTATTAATCACTTAAGTATGATCCATATGG 237
 701 GCACAGTTCAATTTCCGCGGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 760
 238 GAGATACACCAAAAATCTAAACACACGACACTGATAAGT 279
 761 GATGACACATCAAGAAATCAAAAATCTCTGCACTGATTAAGT 802

RESULT 5
 LOCUS BX553516 398 bp mRNA linear EST 10-OCT-2003
 DEFINITION BX553516 Glossina morsitans morsitans adult infected gut Glossina morsitans morsitans cDNA clone Tse129h11.q1c, mRNA sequence.
 ACCESSION BX553516
 VERSION BX553516.1 GI:33377668
 KEYWORDS EST.
 SOURCE Glossina morsitans morsitans
 ORGANISM Glossina morsitans morsitans
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Hippoboscidae; Glossinidae; Glossina.
 1 (bases 1 to 398)
 REFERENCE
 1 Lehane,M.J., Aksoy,S., Gibson,W., Kethornou,A., Berriman,M., Hamilton,J., Soares,M.B., Bonaldo,M.F., Lehane,S. and Hall,N. Adult midgut expressed sequence tags from the tsetse fly *Glossina morsitans morsitans* and expression analysis of putative immune response genes

KEYWORDS	EST
SOURCE	Anopheles gambiae (African malaria mosquito)
ORGANISM	Anopheles gambiae
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anophelinae
REFERENCE	1 (bases 1 to 632)
AUTHORS	Holt, R.A., Lin, Y.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab, R., Collins, F.H., Venter, J.C. and Hoffman, S.L.
TITLE	Celera Anopheles gambiae EST project
JOURNAL	Unpublished (2002)
COMMENT	Contact: Holt R.A. Celera Genomics 45 W. Gude Dr., Rockville, MD 20850, USA Tel.: 240453151 Fax: 2404534580 Email: HoltRA@celera.com Plate: NU01004SFB row: F column: 10 Seq primer: M13 Reverse
FEATURES	Location/Qualifiers
Source	1..632 /organism="Anopheles gambiae" /mol_type="mRNA" /strain="RSP-ST (Reduced susc. to Permethrin - std. chromosome)" /db_xref="taxon:7165" /clone="19600447009427" /dev_stage="Adult" /lab_host="DHI10b" /clone_lib="A.Gam.ad.cDNA1" /note="Vector: pSport1; Site 1: SalI, Site 2: NotI, whole adult mosquitoes (mixed sex) frozen on liquid nitrogen. cDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)."
ORIGIN	
Query Match	13.9%; Score 69.2; DB 4; Length 632;
Best Local Similarity	52.7%; Pred. No. 1.1e-08;
Matches 149; Conservative	0; Mismatches 134; Indels 0; Gaps 0;
QY	194 GATGCGGTGATCTTTATCTACACTTATAGTGATCTCTATATGAGCATACACCAAA 253
Db	350 GTGTGACCTGTGTGTTACTTGTAGGAGTCTGTGTGTGTGTGACCATCCGCCGAC 409
QY	254 AATGTAAACACGACGACATGATTAAGTTTCTTATTTCTTGATGCGGACCTGCA 313
Db	410 AATGAAATTCGACACGACTCTGTGCATACGATCTCCCTGGACAGTTTGCCGCTGG 469
QY	314 CCGGTATGCTTGGCTGACGCTGTCTTAAAGTTCTACATTTGCGCAAGTTGACGCCAA 373
Db	470 ATGATTTTGGCGCTGGATACCTTTGGCGCTGAAGATTTACATGATCTTACGAAGACGCCA 529
QY	374 CAATTGATTAATGATCTTGAAGTGTANAGCATGGTGTGAAGTTAGTTGGGATATGAG 433
Db	530 TCATTGACAGCTCCGATCCACGCGTGTGGAACATGTGTGCTCGCTGATATATGCTCG 589
QY	434 GATCACTTATATTTTGGCAACGTTGATTTGATTTGCTTC 476
Db	590 GTGTGGCGATGATCATCTTGGCTTTCTGATCGGAATGTTCC 632
RESULT 10	
LOCUS	BX551134
DEFINITION	BX551134 Glossina morsitans morsitans adult infected gulo Glossina morsitans morsitans cDNA clone Tcell16908_q1c, mRNA sequence.
ACCESSION	BX551134
VERSION	BX551134.1 GI:33374988
KEYWORDS	EST.
SOURCE	Glossina morsitans morsitans
ORGANISM	Glossina morsitans morsitans

REFERENCE	AUTHORS	TITLE
JOURNAL MEDLINE PUBMED	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preyigota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Hippooscolidae; Glossinidae; Glossina.	1 (bases 1 to 561)
Lehane, M. J., Aksoy, S., Gibson, M. W., Kethornou, A., Berriman, M., Hamilton, J., Soares, M. B., Bonaldo, M. F., Lehane, S. and Hall, N.	Adult midgut expressed sequence tags from the tsetse fly Glossina morsitans morsitans and expression analysis of putative immune response genes	Genome Biol. 4 (10), R63 (2003)
14519198	Contact: Hall N	
Pathogen Sequencing Unit	The Sanger Institute The Wellcome Trust Genome Campus	
Hinxton, Cambridgeshire, CB10 1SA, UK	Request for clones, please contact: Mike Lehane	
Prof. M. J. Lehane	School of Biological Sciences,	
University of Wales,	Bangor LL57 2UW	
All clones with suffix q1c are reverse primer reads starting at 5' end of the cDNA all pic reads are from the 3' end.		
FEATURES	Location/Qualifiers	
Source	1..561	
/organism="Glossina morsitans morsitans"		
/mol_type="mRNA"		
/sub_species="morsitans"		
/db_xref="taxon:37546"		
/clone="Tse116g08 q1c"		
/issue_type="adult infected gut"		
/clone_id="Glossina morsitans morsitans adult infected gut"		
/note="country: Zimbabwe; EST from adult gut infected with T.brucei"		
ORIGIN		
Query Match	13.3%; Score 66; DB 5; Length 561;	
Best Local Similarity	59.4%; Pred. No. 9e-08;	
Matches 111; Conservative 0; Mismatches 76; Indels 0; Gaps 0;		
QY	304 GGACCTGCAACCGTTATGCTTGGCTAGCGCTGCTGAAGTTCTACATTTCCGCAAG 363	
Db	2 GGTCCGATGAGGGCTTTTTCGGTATATTTCTTAAATATGTTATGATCGACT 61	
QY	364 TTGACGCAACATGATTAATATATGATGCTGATGATTAGAGCATGGTGAAGTTGG 423	
Db	62 AAAATCTCTTAAATGATTAATAGATTCGGGTTGGCTGAGACGTGGTGGTTGGTTGG 121	
QY	424 GTGATCTAGGATCACTTAATATATTTTTCGCAAGTTGATTTGATTTGTTCTTAAATT 483	
Db	122 ATGATCTTGGTACTATGATGATATTTATTTTCCGGTTAATAGGCTTATTTCTTAAAGAG 181	
QY	484 CTGCCAA 490	
Db	182 CTACCTA 188	
RESULT 11		
AU209481	463 bp	RNA
AU209481	unpublished oligo-capped cDNA library, stage L1	linear
LOCUS	Caenorhabditis elegans cDNA clone YK1015F10 5', RNA sequence.	EST 17-JUL-2001
DEFINITION		
ACCESSION	AU209481	
VERSION	AU209481.1	GI:14845081
KEYWORDS	EST.	
SOURCE	Caenorhabditis elegans	
ORGANISM	Caenorhabditis elegans	
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Pelodermidae; Caenorhabditis.	
AUTHORS	1 (bases 1 to 463)	
Kohari, Y., Shin, I., T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.		

AUTHORS	Kohara, Y., Shin-i, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.
TITLE	A complementary view of the <i>C.elegans</i> genome
JOURNAL	Unpublished (2001)
COMMENT	On Jul 17, 2001 this sequence version replaced gi:14832514.

FEATURES	Location/Qualifiers
source	1. .484

0342

Query Match	8.6%	Score 43	DB 1	Length 484
Best Local Similarity	46.8%	Pred. No.	0.35	
Matches 133	Conservative 0	Mismatches 151	Indels 0	Gaps 0

[illegible]

RESULT 14				
CL65155				
CL65155	840 bp	DNA	linear	GSS 25-AUG-2004
t2c08be.t1	TAM8T Bos taurus genomic clone	t2c08be,	genomic survey	sequence.

REFERENCE	1. (bases 1 to 840)
AUTHORS	Lin, S., Natar, P. Z., Adelson, D., Gill, C. A. and Roe, B. A.
TITLE	Bovine BAC End Sequences from Library TAMBET
JOURNAL	Unpublished (2003)
COMMENT	Contact: Bruce A. Roe

Fax: 405 325 7762
Email: broe@ou.edu
Class: BAC ends
High quality sequence start: 396
High quality sequence stop: 654

FEATURES
SOURCE

ORIGIN

Query Match	8.6%	Score 42.8;	DB 9;	length 840;
Best Local Similarity	50.0%;	Pred. No. 0.45;		
Matches 107; Conservative	0;	Mismatches 107;	Indels 0;	Gaps 0;

QY	50	AGCATCAAAATGTCACCACTTGAAGTCTTAATTAACGAAAGGAAACCTTGGCCACA	109
Db	422	AGGAACAGACGCGCTGCCCTTAAATGTAAACTTAACATTACAGGAATCTTTGCCATAAA	481
QY	110	CTGATGCGGAGAGATTGAATCGGAACAGAAAGGAGATGCGACCTCAAAATATATAC	169
Db	482	CTAAAAAGCTGTTTGACCTMAAGCAACATCTTTGGCAGATGACAGAGTTGTATGA	541
QY	170	TATTTGCTGCCAATTAATTTCTGAGCGGTGATCTTTATCTACACTTTAGGTAT	229
Db	542	GCCTGAGACTTAAATTACTTCTTTGCACTAATTAAGGGGACTCTCAAGGTACAGCTTAT	601
QY	230	CCTATATGAGCATATACACCAAAAATCTAATAAC	263
Db	602	GGCATATTGATGGAGATTAAGAAATCTTTACAC	635

RESULT 15	CO715174/c	LOCUS	DEFINITION
	CO715174	555 bp	mRNA linear
	CG14-24518	DG14-muscle	Canis familiaris CDNA 3', mRNA sequence.

REFERENCE	1 (bases 1 to 555)
AUTHORS	Schlueter, T., Hermanns, J., Weindel, M., Schuetz, D., Kranz, H., Henrich, J., and Joebett, R.
TITLE	Dog arrayTag cDNA clone collection
JOURNAL	Unpublished (2004)
COMMENT	Contact: Thomas Schlueter LION bioscience AG Walhoferstrasse 98, D-69123 Heidelberg, Germany Tel: +49 6221 4038 150 Fax: +49 6221 4038 290 Email: Thomas.Schlueter@lionbioscience.com, Location/Qualifiers 1..555
FEATURES	
source	1..555

```
FEATURES
source
location/Qualifiers
1..555
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="Beagle"
/db_xref="taxon:9615"
/tissue_type="muscle"
/dev_stage="adult"
/lab_note="DH10B"
/clone_id="Dg14-muscle"
```


